FACIAL EXPRESSIVITY IN PARKINSON'S DISEASE:

USING THE FACIAL ACTION CODING SYSTEM TO EVALUATE DUCHENNE SMILING BEHAVIOR AND THE IMPACT OF VOICE TREATMENT UPON GLOBAL EXPRESSIVITY MEASURES AT 6-MONTH FOLLOW-UP

by

DAVID MCCABE

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the degreeof Doctor of Philosophy, The City University of New York.
2013

Print Form

The manuscript has been read and accepted for the
Graduate Faculty in Psychology in satisfaction of the
Dissertation requirements for the degree of Doctor of Philosophy

(Print) Joan C. Borod, Ph.D

(Signature)
Chair of Examining Committee

Date

(Print) Maureen O'Connor, Ph.D.

(Signature)
Executive Officer

Date

(Print) Laura Rabin, Ph.D

(Print) Justin Storbeck, Ph.D

(Print) Yvette Caro, Ph.D

(Print) Deborah Walder, Ph.D.

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

FACIAL EXPRESSIVITY IN PARKINSON'S DISEASE:
USING THE FACIAL ACTION CODING SYSTEM TO EVALUATE DUCHENNE SMILING BEHAVIOR AND THE IMPACT OF VOICE TREATMENT UPON GLOBAL EXPRESSIVITY MEASURES AT 6-MONTH FOLLOW-UP

by

David McCabe

Dissertation Chairperson: Joan C. Borod, Ph.D.

Nonverbal signals contribute significantly to interpersonal communication. Facial expressivity, a major source of nonverbal information, can be compromised in Parkinson's disease (PD). The resulting disconnect between subjective feeling and objective facial affect can lead people to form negative and inaccurate impressions of people with PD with respect to their personality and intelligence. Previous research (Spielman, Borod, & Ramig, 2003) suggests that the Lee Silverman Voice Treatment (LSVT LOUD; Ramig, Pawlas, et al., 1995) might benefit facial expressivity in PD. To better understand the nature and psychosocial impact of facial expression deficits, a two-component study was conducted. First, the long-term (6-month) efficacy of LSVT LOUD was compared to a second intervention (ARTIC), which targets articulation, in treating facial expressivity changes in PD. Global measures of facial expressivity were used to study 6-month follow-up data and build upon the pre-/post- findings of Dumer (2011). Second, smile behavior was examined at baseline and as a function of treatment condition and time. Smile frequency, intensity, and onset duration data were examined, and Duchenne smiles, commonly thought to reflect spontaneous or "felt" emotion, were

distinguished from non-Duchenne smiles. Data were obtained from video footage of healthy controls (age matched; *n* = 11) and individuals with PD (*n* = 45). The PD group was comprised of individuals receiving no treatment (*n* = 17), individuals receiving an articulation-based treatment (Artic; *n* = 12), and individuals receiving LSVT LOUD (*n* = 16). Video footage was obtained at baseline, post-intervention, and 6-month follow-up. Facial expressions were coded using the Facial Action Coding System (FACS) developed by Ekman and Friesen (1978). At baseline, Healthy Controls generally exhibited higher levels of facial expressivity as compared to individuals with PD, though gender effects may have contributed to these findings. At 6-month follow-up, global measures of facial expressivity did not significantly differ across treatment groups. Although the LSVT group increased in some measures of smile behavior, LSVT did not generally differ from other treatment conditions in degree of treatment impact over time, as assessed by a nonparametric analysis of change-scores.

ACKNOWLEDGMENTS

This manuscript is the culmination of much planning and perseverance. It was, from start to finish, a team effort.

First and foremost, I would like to thank Dr. Joan C. Borod, chair of my dissertation committee, research advisor, and mentor. Joan has provided much guidance and support over the years, and for this, I am very grateful. I must also thank my dissertation committee. Dr. Oster's methodological expertise was invaluable in developing this research proposal. Dr. Rabin's feedback has been consistently insightful and constructive—cogency and clarity have increased under her watch. Dr. Storbeck's analytical prowess and encouragement were great assets during the final stages of the project. And of course, I am indebted to the readers, Dr. Yvette Caro and Dr. Deborah Walder—it is the kind soul and committed professional who adds a dissertation to their summer reading list. To all of these individuals, I extend my heartfelt thanks.

I must also acknowledge the enormous contribution of Dr. Aleksey Dumer. My dissertation builds upon Aleksey's doctoral research. Aleksey is a brilliant mind, a colleague of the utmost professionalism, and was a fantastic labmate. I would also like to thank Dr. Kim Savage. As I was just beginning to cut my teeth in graduate school, Kim was a more than capable ally in navigating SPSS, the IRB, or any other acronym to be found in the academic sphere. The Emotion Lab—past and present—is a good group. Their expertise, humor, and support are much appreciated.

This project was supported by NIH grant R01-DC01150, awarded to Dr. Lorraine O. Ramig at the University of Colorado at Boulderand the National Center for Voice and Speech. Dr. Ramig's work on LSVT was foundational to this project. Jennifer Spielman, also in

Colorado deserves honorable mention. Jennifer was extremely helpful in ensuring that the inter-site collaboration ran smoothly, and her published insights on LSVT were a valuable reference in formulating this research.

Finally, I would like to thank my family and friends. They have supported me in myriad ways.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

Introduction

The human face is widely considered to be the primary vehicle of emotional expression (e.g., Borod &Koff, 1984). Through various configurations of facial muscle movements, as well as head and eye movements, we convey information—emotional and non-emotional, spontaneous and deliberate—to express (or mask) our feelings and beliefs, and even to visually "punctuate" (e.g., Ekman & Friesen, 1975, p. 119) spoken language. In Parkinson's disease (PD) and other neurological conditions, this communication channel can be compromised.

A number of potential treatments for PD-related facial expressivity deficits have been explored. One such intervention is the Lee Silverman Voice Treatment (LSVT®LOUD; Ramig, Pawlas, & Countryman, 1995: Ramig et al., 2001). Although developed to treat speech changes in PD, limited data (Spielman et al., 2003) suggest a possible effect on facial expressivity as well. The proposed study has two broad goals: 1.) to compare the long-term (6 month) efficacy of LSVT LOUD and an articulation-based voice treatment on facial expression in PD; and 2.) to understand the impact of these interventions on smiling behavior through an examination of Duchenne and non-Duchenne smiles.

The value of establishing facial-expressivity treatment efficacy at 6 months post-treatment is straightforward. With access to follow-up data, clinicians will be better able to estimate treatment effects, inform patients about potential benefits, and use follow-up data to optimally space "booster" treatment sessions, if clinically indicated.

Through an investigation of smiling behavior in PD, we hope to understand better the impact of disease and treatment upon this psychosocially important class of facial expression. Duchenne and non-Duchenne smiles are thought to reflect two different relationships between face and emotion. Duchenne smiles, which are characterized by the combined effect of

upturned/stretched lips as well as contraction around the eye, are frequently associated with "felt" emotion. In contrast, non-Duchenne smiles, which are characterized by upturned lips in the absence of eye involvement, appear to be less directly tied to subjective pleasant emotion (e.g., they may be produced deliberately as a social nicety; e.g., Ekman, Davidson, & Friesen, 1990).

The following review outlines the anatomical factors involved in facial expression, considers the possible etiology of facial expression deficits in PD, explores the significance of changes in Duchenne smiling in PD, and examines behaviorally-based treatment options for PD-related deficits in facial expression.

The Human Face

The human face can be broadly divided into three distinct regions. Based on their largely distinct motoric connectivity and contributions to the facial message (e.g., lower face – speech; upper face – punctuation of speech), these regions are: brows/forehead, eyelids/root of nose, and lower face (which includes the cheeks, mouth, lower nose, and chin; Ekman & Friesen, 1975). In almost all individuals, the main branch of the facial nerve divides into an upper (temporofacial) and a lower (cervicofacial) portion (Rinn, 1984). The cervicofacial division supplies the lower face. The temporofacial division gives rise to the zygomatic and temporal branches, which, respectively, innervate the middle and upper face.

The temporofacial division typically conveys signals generated by either hemisphere. The proportion of contralateral direct corticobulbar fibers innervating the orbicularis oculi (located in the upper face) is estimated to be about 75% (with the remaining 25% being ipsilateral in origin). The extent of bilateral innervation in the brow and forehead region of the upper face is even greater, with approximately equal degrees of ipsilateral and contralateral innervation (Borod &Koff, 1984; Rinn, 1984).

By contrast, the cervicofacial division, which innervates the lower face, typically transmits signals originating from the contralateral hemisphere. The direct corticobulbar fibers that innervate the lower half of the face are commonly described as originating predominantly or exclusively from the contralateral hemisphere (Borod &Koff, 1984; Borod, Haywood, &Koff, 1997; Rinn, 1984; c.f., Morecraft et al., 2001). Although this description captures typical organization, it is important to note that individuals show variability both in the path of the peripheral portions of the facial nerve and in the correspondence between the particular branch and the muscles that they ultimately innervate (Rinn, 1984). Also, the relative degree of ipsilateral versus contralateral innervation to the lower and upper face remains a matter of debate (for discussion, see Morecraft et al., 2001).

In discussing facial motor control, it is important to distinguish between voluntary (i.e., deliberate) movements and spontaneous facial expression. To a large extent, deliberate and spontaneous facial expressions rely upon distinct, non-overlapping neuroanatomical pathways, with deliberate facial movement dependent upon the pyramidal tract and spontaneous facial movement dependent upon the extrapyramidal tract, which originates from premotor cortical areas and subcortical regions (e.g., basal ganglia; Borod, 1993a; Borod &Koff, 1984; Borod, Lorch, Koff, & Nicholas, 1987; Monrad-Krohn, 1924; Rinn, 1984). With respect to voluntary movement, the left and right lower hemifaces are largely functionally independent of each other as evidenced by the observation that individuals can perform unilateral (i.e., hemifacial) lower face movements. The left and right forehead musculature, however, have been described as "neurologically yoked" (Rinn, 1984, p. 58) such that unilateral muscle contraction of the upper face is more difficult.

Notably, some evidence suggests that the functional independence of these pathways is not total (e.g., Jox, Bruning, Hamann, &Danek, 2004). For a more recent review of the neuroanatomy of facial expression, which addresses differences between upper and lower face connectivity, including differences between volitional and spontaneous emotional expression, see Morecraft, Stilwell-Morecraft, and Rossing (2004).

Parkinson's Disease: Background, Symptomatology, and Neuropathology

Overview

To establish the social, medical, and psychological context in which Parkinsonianmasked facies arises, it is necessary to briefly examine the demographic distribution of PD, review the neuropathologic features of the disease process, and highlight both motor and non-motor clinical manifestations.

Although the term "masked facies" has yet to enter lay vocabulary, PD is a relatively common, age-related, neurodegenerative disease afflicting 1-2% of people aged 60 yearsand older. Because the elderly population in the U.S. is growing (CensusScope, 2009 [based on the 2000 U.S. census]), the frequency of PD in the U.S. is expected to increase dramatically over the coming years. Nevertheless, PD afflicts a small enough portion of the population that the typical layperson will have a very limited familiarity with PD symptomatology. It is reasonable to assume that limited familiarity with PD-related communication challenges (including but not limited to masked facies)—on the part of the patient, the patient's family and friends, the general public, and various clinicians—contributes to the psychosocial impact of living with PD.

Parkinson's disease, also known as primary parkinsonism (Fahn, 2003) or idiopathic PD, is the most common form of parkinsonism (for review, see Olanow, Stern, &Sethi, 2009). Cardinal features of PD include resting tremor, bradykinesia, rigidity, gait disturbance, and

postural instability. Diagnostic accuracy is extremely high for patients exhibiting resting tremor, prominent asymmetry of disease symptoms, and responsiveness to levodopa (Olanow et al., 2009). Additional motor features include micrographia, freezing, flexed posture, and changes in facial expressivity, such as masked facies and decreased blinking (Olanow et al., 2009). Non-motor symptoms (e.g., depression, anxiety disorders, abulia, aspontaneity, apathy, cognitive deficits, autonomic dysfunction, pain, sleep disturbances, and hallucinations/illusions/delusions) are also present in PD, and recent investigations have provided a better understanding of the extent and nature of this aspect of PD pathology (for reviews, see Olanow et al., 2009; Zgaljardic, Borod, Foldi, &Mattis, 2003, 2006; Zgaljardic, Foldi, & Borod, 2004). Voice and speech deficits, which result from a combination of motor and sensory-perceptual changes, are also present in PD (for review, see Fox, Morrison, Ramig, & Sapir, 2002) and are discussed in greater detail within the treatment subsection of this research proposal.

Depression and other Psychiatric Features in PD

The non-motor symptoms of PD affecting mood and emotional processing are of particular relevance to masked facies (for review, see McCabe, Borod, Meltzer, Spielman, & Ramig, 2010). Depression, perhaps the most relevant emotional processing disorder with respect to diminished facial expressivity, is experienced by approximately 40% of PD patients at least once during the course of their illness (Cummings, 1992; Frisina, Borod, Foldi, &Tenenbaum, 2008; Mayeux, Stern, Rosen, &Leventhal, 1981; Raskin, Borod, & Tweedy, 1990; Sano et al., 1989; Starkstein, Preziosi, Bolduc, & Robinson, 1990; Zgaljardic et al., 2003). The origin of depression in PD may be reactive (i.e., an emotional response to stresses of diagnosis, treatment, and disease progression) or may arise directly from PD pathology (e.g., as a result of decreased brainstem dopaminergic, noradrenergic, and/or serotonergic neurons [Cummings &Masterman,

1999; Ravina, Edwards, & Sheehy, 2003]). Of note, the frequent occurrence of both motor symptoms (including masked facies) and depression in PD is clinically challenging. Clinicians may mistakenly diagnose depression based on non-psychiatric symptoms (including masked facies) or, more likely, may underdiagnose depression in PD (Shulman, Taback, Rabenstein, & Weiner, 2002) as a result of attributing somatic depressive symptoms (such as loss of appetite, loss of libido, insomnia, etc.) to somatic features due to PD pathology as opposed to psychiatric/psychological origins. It is estimated that only 20% of individuals with comorbid depression and PD are actually treated for depression (Frisina, Borod, et al., 2008).

Although depression is probably the most commonly discussed mood correlate of PD, it is certainly not the only psychiatric disturbance associated with the disease. Anxiety disorders alone, or in combination with depression, are experienced by about 40% of individuals with PD and, like depression, may be a reaction to living with PD or may stem more directly from PD-related pathology (Siemers, Shekhar, Quaid, & Dickson, 1993; Stein, Heuser, Juncos, &Uhde, 1990). Apathy, not necessarily co-occurring with depression, is also frequently seen in PD (for review see Zgaljardic et al., 2007). Notably, although apathy clearly impacts emotional experience, some researchers underscore the cognitive as opposed to emotional components of this disorder (Robert et al., 2006; Starkstein et al., 1992). Finally, other mood disturbances associated with PD (e.g., pseudobulbar affect) may even present with increased emotionality (Cummings et al., 2006).

Neuropathology

Overview of Neuropathology in PD. The pathological hallmarks of PD include a loss of neuromelanin-containing monoamine neurons (dopamine [DA], in particular) in the substantianigra pars compacta as well as the presence of Lewy bodies in monoamine neurons

(Fahn, 2003). Non-dopaminergic pathology is seen in cholinergic neurons of the nucleus basalis of Meynert, norepinephrine neurons of the locus coeruleus, and serotonin neurons of the midline raphe, as well as in neurons within the cortex, brainstem, spinal cord, and peripheral nervous system (Olanow et al., 2009).

Emotional Processing Deficits in PD. While the neuropathology of global motor deficits in PD is clearly relevant to a discussion of facial expressivity, it is also necessary to appreciate the capacity of PD to impact the neural substrates of the primary antecedents of facial affect, that is, emotional processing and mood. Multiple emotional processing changes have been observed in PD, and multiple PD-related neuroanatomical factors have been implicated in these changes. The basal ganglia (BG), in particular, are known to play a major role in the experience and communication of emotion (Borod, 2000), and BG degeneration is a signature feature of PD pathology. The limbic system, characterized by intimate connections with the BG and implicated in various facets of emotional processing (MacLean, 1949), is also impacted in PD. Dysfunction of anterior cingulate circuitry, for example, may account for the apathy observed in PD (Tekin& Cummings, 2002; Zgaljardic et al., 2003). The amygdala—a structure whose dysfunction is associated with depression (for a review of functional neuroimaging data, see Drevets, 1998), anxiety, and abnormal fear processing (for review, see LeDoux, 2000)—is innervated by both DA and noradrenergic systems (Fallon, Koziell, & Moore, 1978; Fudge & Emiliano, 2003), both of which are compromised in PD. Post-mortem studies in PD have revealed volume reductions in the amygdala (up to 20%) and Lewy body pathology within the BG (Harding, Stimson, Henderson, &Halliday, 2002).

Of the various psychiatric disorders occurring in PD that might be relevant to masked facies, depression has probably been the most extensively researched. Depression in PD has

been related to a number of anatomical and neurophysiological changes, including low levels of 5-HIAA (Mayeux et al., 1986), dysfunction of the mesocorticolimbic DA system (Mayberg & Solomon, 1995), and dysfunction of frontal lobe circuitry (Mayberg et al., 1990). Some argue that the role of DA system pathology is limited, or not directly causal, with respect to depression in PD and that the evidence for serotonergic dysfunction is more compelling (e.g., Sano, Stern, Cote, Williams, &Mayeux, 1990). This view is supported by the mixed evidence regarding the efficacy of DA therapy in treating depression in PD in contrast to demonstrated success using tricyclic antidepressants and SSRIs (Frisina, Tenenbaum, Borod, & Foldi, 2008; Kulisevsky et al., 2008; Zgaljardic et al., 2003). Distinguishing between these systems is difficult as dysfunction of DA pathways may have an indirect impact on the serotonergic system (Zgaljardic et al., 2003). Also of clinical interest, high frequency deep brain stimulation of the subthalamic nucleus has resulted in improvements in subjective well being and appears to also have antidepressive effects (e.g., Schneider et al., 2003). On a conceptual note, the psychiatric entities resulting from PD pathology might not necessarily fit with well-established diagnostic classifications. Some investigators argue that, in PD, symptoms that are often classified by clinicians as depression may in fact be more accurately viewed as a distinct affective disorder arising from PD-related caudate pathology (Taylor & Saint-Cyr, 1990).

Emotional and Motoric Contributions to Facial Expressivity Changes in PD

Facial Expression in PD

Although masked facies (also commonly referred to as “hypomimia”) is present in almost all cases of PD (Hoehn &Yahr, 1967; Jankovic, 2003), it is not considered to be a cardinal feature of the disease and can be symptomatic of a number of disorders. Thus, masked facies is

frequent in, but not specific to, PD (e.g., Galasko, Katzman, Salmon, & Hansen, 1996; Hohl, Tiraboschi, Hansen, Thal, & Corey-Bloom, 2000).

Masked facies can be interpreted as an extrapyramidal sign but, in the context of PD, multiple factors can contribute to "masked" or diminished facial affect. Concerning the contribution of motoric factors, Stochl and colleagues (Stochl, Boomsma, Ruzicka, Brozova, &Blahus, 2008) employed structural equation modeling and found that "speech/hypomimia [i.e., masked facies]" represented one of five main latent symptom factors on the Motor section of the Unified Parkinson's Disease Rating Scale (UPDRS; Fahn& Elton, 1987) and was correlated (.54 to.85) with other factors relating to rigidity and bradykinesia. However, in addition to its presence in progressive neurological conditions, diminished facial expressivity is also symptomatic of depression (e.g., Mergl, Mavrogiorgou, Heger, & Juckel, 2005), other psychiatric conditions, and certain pharmacological side-effects (e.g., Juckel et al., 2008). Thus, while PD-related motor deficits likely contribute to masked facies, the contribution of emotional processing deficits and mood disorders, arising directly or indirectly from PD, must also be considered.

Emotion and Expression in PD

From Perception to Expression: Emotional Processing Deficits in PD. Emotional processing involves a multitude of interconnected subprocesses including perception of environmental/social stimuli, valence appraisal, physiological arousal, subjective experience, and expression (for conceptualizations of emotional processing components, see Borod [1993b] and Heilman, Blonder, Bowers, and Valenstein [2003]). Broadly speaking, emotional expression can be viewed as the (possible) endpoint of emotional experience. Thus, theoretically, diminished facial expressivity may arise at stages of processing affecting the periphery (e.g., resulting from a

compromised facial nerve) or may follow from the dysfunction of any number of more central, antecedent, processes (e.g., emotional perception, appraisal, etc.). See Figure 1.

At what stage in the sequence of events culminating in facial affect does PD compromise emotional expression? Is masked facies purely an expressive deficit, per se, resulting from a global motoric deficit? Or, is masked facies, at least in part, the intact expression of diminished or altered emotional experience? To better understand the nature of masked facies, specifically, in PD, it is necessary to examine the broader pattern of emotional functioning across multiple communication channels and multiple modes (i.e., levels or stages) of emotional processing. Communication channels include face, prosody, speech content, gesture, posture, and scenes (a "scene" refers to the environment or a pictorial representation in which an emotional situation occurs; Borod, Tabert, Santschi, & Strauss, 2000). Emotional processing modes include perception, experience, and expression (Borod, 1993b).

In a review of the emotional processing deficit literature in PD, the Emotion Lab at Queens College identified 23 relevant studies (Borod & Brickman, 2001; Zgaljardic et al., 2003), involving 35 experimental observations (i.e., research findings). In terms of emotional processing mode, 55% of the observations involved expression; 39%, perception; and 6%, experience. In terms of communication channel, 54% of the observations involved faces, and the remaining observations involved prosody, gesture, speech content, posture, or scenes.

The review (Borod & Brickman, 2001; Zgaljardic et al., 2003) revealed that the research findings are not in agreement regarding the nature of PD deficits in emotional processing. In terms of emotional perception, findings for significant emotional processing deficits (EPDs) in individuals with PD compared to healthy normal controls were equivocal, with 50% of the facial channel observations revealing deficits, 55% of those for prosody, 0% of those for speech

content, and 50% of those for scenes. Findings for emotional expression were more consistent and generally suggested deficits: 100% of the observations for prosody revealed deficits, 100% for gesture, 100% for posture, and 57% for face. Finally, for emotional experience, in 100% of the studies, PD participants did not differ from healthy controls. Thus, this initial review of the PD emotional processing literature indicated significant deficits in expression, equivocal findings for perception, and no apparent experiential deficits.

Since that initial review, the Emotion Lab conducted a further analysis of 18 studies published between 2000-2006. These studies yielded 22 observations. In terms of processing mode, 32% of the observations involved expression, 64% involved perception, and 4% involved experience. In terms of channel, 57% involved faces; 24%, prosody; 14%, speech content; and 5% scenes. The evidence for significant perceptual EPDs in PD continues to be equivocal, with 57% of the observations for face showing deficits (Dujardin et al., 2004; Glozman et al., 2003; Kan et al., 2002; Sprengelmeyer et al., 2003), 50% for prosody (Breitenstein et al., 2001; Pell & Leonard, 2003), 50% for speech content (Sartorio et al., 2005), and 100% for scenes (Wieser et al., 2006). In a recent meta-analysis of studies examining emotion recognition/perception deficits in PD, Gray and Tickle-Degnen (2010) reported an overall effect size of $g = 0.52$, with negative emotions being particularly affected; neither visuospatial impairment nor depression appeared to account for these findings but the authors highlighted the potential for working memory deficits to negatively impact prosodic emotion recognition. For expression, significant EPDs in PD as compared to healthy controls occurred for 60% of the observations for face (Reid, 2000; Simons, Ellgring, &Pasqualini, 2003; Simons, Pasqualini, Reddy, & Wood, 2004) and for 100% of the observations for prosody (Pell et al., 2006). For the lexical channel, individuals with PD differed from controls but in the opposite direction, producing discourse with greater

emotional intensity (Halfacre, Borod, Pick, Krch, & Gruber, 2006) and longer duration (Crucian, et al., 2001). For experience, there were no significant deficits in the one study examined (Simons et al., 2004). Thus, there continues to be a lack of consensus regarding EPDs in PD.

Further Complexities. Reviewing the PD-related emotional processing literature in terms of perception, experience, and expression, allows one to approximate the stages at which PD-related deficits might contribute to diminished facial affect. More precise identification and definition of PD-related EPDs, however, requires further analysis. Studies have shown that although PD and control participants report comparable valence (pleasantness/unpleasantness) ratings of emotional pictures, PD participants, as compared to controls, appear to experience these pictures as less emotionally arousing (Bowers, Miller, Mikos, et al., 2006; Miller, Okun, Marsiske, Fennell, & Bowers, 2009). Evidencing this possibility, Bowers and colleagues (Bowers, Miller, Mikos, et al., 2006), observed that, despite possessing apparently intact eyeblink motor circuitry, individuals with PD failed to show normal emotion-mediated modulation of the startle eyeblink reflex. So, even in situations in which individuals with PD provide normal valence ratings of a stimulus and are capable of producing normal facial affect, they might exhibit lower reactivity to arousing stimuli and thus be less likely to spontaneously execute observable facial affect. Further complicating the picture, emotional changes accompanying PD are not exclusively characterized by the reduction or attenuation of emotional experience and/or expression. For example, the lexical content of emotional monologues produced by individuals with PD has actually been rated as more emotionally intense than monologues produced by healthy demographically matched controls (Halfacre et al., 2006). Whether this observation reflects a qualitatively different emotional experience for individuals

with PD or a compensatory behavior (e.g., deliberately using emotionally charged words to compensate for reduced non-verbal emotional expressivity) is not known.

Depression, PD, and Masked Facies. In individuals with PD, it is particularly difficult to assess the contribution of depression to masked facies. Depression is common in PD and may be either "reactive" (a psychological reaction to the experience of living with PD) or "organic" (a biologically driven, PD-related depression). Depression, independent of PD-related motor deficits, can contribute to diminished facial expressivity (Bleuler, 1983; Marsh, 2008).

Because individuals with PD and depression can have similar clinical presentations with respect to facial expressivity and because of the responsiveness of both groups to electroconvulsive therapy and tricyclic antidepressants, it has been suggested (Serby, 1980) that PD and depression may share a common underlying neuropathology. However, although diminished facial expressivity is present in both PD and depression, the profile of facial expressivity deficits may differ qualitatively or quantitatively between the two groups. In depressed individuals, facial expression appears to be selectively altered for positive, relative to negative, emotion (Jaeger, Borod, &Peselow, 1986; for review, see Yecker et al., 1999). Individuals with PD, however, as well as healthy controls, have been found to express positive facial emotions more intensely than individuals with unipolar depression, schizophrenia, or right-brain-damage (Borod et al., 1990).Borod and colleagues (1990) also found that PD participants were significantly less accurate than unipolar depressed and healthy control participants in their attempts to voluntarily produce emotional facial expressions.

A study by Katsikitis and Pilowsky (1991) provides additional relevant data. Using digital image analysis of facial landmark changes, these authors measured the smiling behavior of individuals with PD, individuals with major depression, and healthy controls. All participants

viewed amusing slides, and their most pronounced smiles in reaction to the slides were entered into analysis. Facial expressivity was operationalized as a number of objective variables, such as degree of mouth-opening and mouth-width. Both the depression group and the PD group smiled significantly less frequently than the control group. As would be expected, the major depression group obtained higher depression scores than the PD group on a depression inventory developed by the authors. Similarly, the PD group endorsed greater levels of depression than did the control group.

Two findings from the Katsikitis and Pilowsky study (1991) are revealing. Across the groups, depression scores were correlated with multiple facial expressivity variables but when calculated separately for each group, this correlation was not present in either the PD or healthy control group. Additionally, the depressed group exhibited a distinct profile relative to PD and control participants, marked by significantly higher scores on multiple facial landmark measures, "reflecting what is often referred to colloquially as a 'long face'" (Katsikitis&Pilowsky, 1991, p. 687). As the authors note, it is unclear whether this difference between the PD and the depressed group is qualitative or quantitative in nature. Although some investigators highlight the similarity between facial behaviors in PD and depression (citing the bradykinetic features of facial behavior common to both populations; e.g., Mergl et al., 2005), the findings of Katsikitis and Pilowsky (1991) are compatible with the notion that, when comorbid with PD, depression may account for some, but not all, of diminished facial expressivity.

Deliberate versus Spontaneous Facial Expression in PD

As previously noted, deliberate and spontaneous facial expressions rely upon relatively distinct neuroanatomical circuitry, with deliberate expression associated with pyramidal pathways and spontaneous expression dependent upon extrapyramidal pathways. Because of the

global motor deficits associated with PD, one might expect both deliberate and spontaneous facial expression to be compromised. However, based on the overall evidence, it seems that spontaneous facial expression, compared to posed facial expression, is more negatively impacted by PD. For a more detailed review of differences in spontaneous versus posed facial expression in PD, see Halfacre and colleagues (2009). Presumably, this finding is accounted for by the vulnerability of the extrapyramidal pathway and/or the basal ganglia to PD pathology.

Spontaneous Facial Expression. The observations of Monrad-Krohn (1924) are among the earliest indications that PD might be especially detrimental to the expression of spontaneous emotion. More recent evidence – comprised of both impressionistic rating data (i.e., procedures requiring raters to report their impressions of a participant's facial *affect*), as well as relatively more objective anatomically focused observations (i.e., observations about the frequency and intensity of facial *muscle contraction/relaxation*) – has generally supported this early observation. In one of the early controlled experimental studies of spontaneous facial expression in PD (Buck & Duffy, 1980), individuals with PD were observed as they reacted to emotionally evocative stimuli and were found to be less emotionally expressive than controls. Furthermore, raters had difficulty inferring the content of the emotionally evocative stimuli based upon their observations of the PD participants' facial expressions (Buck & Duffy, 1980). In a more recent study using a range of impressionistic measures to evaluate facial expressions produced during monologues about emotional experiences, Borod and colleagues (2007) reported that individuals with PD were generally less expressive than demographically matched healthy controls. In studies using an anatomically based system (The Facial Action Coding System [FACS] developed by Ekman and Friesen [1978]) to code facial expressivity in PD, investigators found that, compared to controls, individuals with PD exhibited reduced overall expressivity (i.e., fewer facial

movements per evocative scene; Smith, Smith, & Ellgring, 1996), reduced mobility (i.e., expressions were comprised of fewer facial movements; Simons et al., 2003), decreased intensity during smiles (Smith et al., 1996), and fewer Duchenne smiles (Simons et al., 2003; Smith et al., 1996). Similarly, using the Maximally Discriminative Facial Movement Coding System (Max; Izard, 1983), Brozgold, Borod, Rosen, and Alpert (1999) found that individuals with PD exhibited less facial expressivity than controls when engaged in spontaneous displays of facial emotion but did not appear to differ from controls during posed/voluntary conditions. Furthermore, a study that compared spontaneous and posed facial expressions in the same PD population (using the FACS coding system) found that spontaneous smiles, relative to posed smiles, of PD participants with severe PD symptoms were diminished relative to those produced by healthy aged-matched controls and PD participants with milder disease symptoms (Smith, Smith, & Ellgring, 1996).

Posed Facial Expression. Impressionistic and anatomically based observations have also been employed in the evaluation of posed facial expression in PD. Most findings indicate that individuals with PD exhibit reduced intensity of posed emotional expressions relative to controls (Borod et al., 1989; Jacobs et al., 1995; Simons et al., 2004). Additionally, using impressionistic ratings, most investigators have found that raters are better able to identify the intended posed emotions of healthy control participants than those produced by PD participants (Borod et al., 1990; Jacobs, Shuren, Bowers, &Heilman, 1995; Madeley, Ellis, &Mindham, 1995; Simons et al., 2004; for review, see Heilman, Blonder, Bowers, &Crucian, 2000). Studies employing both impressionistic ratings and FACS methodology have also revealed that individuals with PD, as compared to controls, are less able to imitate facial expressions (Simons et al., 2003, 2004). Using a computerized face processing method, Bowers, Miller, Bosch, and

colleagues (2006) found that posed expressions in PD exhibit bradykinesia and hypokinesia (decreased total movement)—the same types of motor deficits that characterize other non-facial intentional movements in PD.

The Relationship between Deliberate and Spontaneous Facial Expression. In natural conversation, facial communication is a combination of spontaneous and posed actions. Posed/deliberate facial actions play an important modulatory role in accentuating, simulating, and "masking" (i.e., intentionally suppressing) spontaneous output, and ultimately provide us with a means to present ourselves in ways that are in accord with social display rules (Ekman & Friesen, 1969). Complementing the investigations of experimentally isolated spontaneous/posed facial expressions discussed above, Simons et al. (2004), using impressionistic ratings, observed that PD individuals also exhibit decreased facial mobility on masking tasks when compared to healthy controls.

The Impact of PD on Discrete Emotions

In addition to observed differences in deliberate/posed and spontaneous emotional expressions in PD, there is also evidence that PD selectively impacts the expression of some emotions more than others. In the context of posed facial expression, individuals with PD, compared to control participants, appear to be especially impaired in producing expressions of happiness, surprise, and disgust (Madeley et al., 1995; Simons et al., 2003 & 2004). Studies of spontaneous emotion have revealed a similar pattern of performance, with PD participants showing impairments in the expression of happiness and disgust (Smith et al., 1996). Halfacre et al. (2009), however, note the possibility that a portion of the expressions categorized as spontaneous in the Smith et al. study might have been deliberately executed in response to social display rules.

The reason for these apparently selective deficits in emotional expression is not known. One possibility is that, compared to other emotions, the expressions of happiness, disgust, and surprise are, from a motoric perspective (e.g., in terms of muscle contraction and muscle coordination), more difficult to execute. Accordingly, what appear as selective emotional impairments, per se, may actually be artifacts of a generalized motoric expression deficit. However, as noted by Halfacre and colleagues (2009), another possibility is that the disease burden in PD is strategically distributed such that some emotional expression pathways are more affected than others (i.e., basal ganglia pathology may selectively impair happiness and disgust).

The Duchenne Smile

Overview and Background

Although a number of facial expressions are negatively impacted by PD, one expression of particular interest is smiling—especially the Duchenne smile (Duchenne, 1862/1990). The Duchenne smile is characterized by the activation of two facial muscle groups: zygomatic major and orbicularis oculi. Whereas a non-Duchenne smile recruits only the zygomatic major muscles (resulting in stretched/upturned lip corners), the added recruitment of the orbicularis oculi in the Duchenne smile results in an added observable effect, that is, wrinkling at the outside corner of the eyes (i.e., "crow's feet") and raising of the cheeks (Duchenne, 1862/1990; Ekman, 1989).

The Duchenne smile is of empirical interest for a number of reasons. First, it is thought to provide some indication concerning whether a smile is "felt" or "unfelt." The presence of a Duchenne smile, compared to a non-Duchenne smile, tends to be associated with the experience of positive emotion (e.g., Ekman, Davidson, & Friesen, 1990). Evidencing this relationship, investigators have found that, in healthy individuals, watching a pleasant film (in both social and solitary viewing conditions), as well as self-reports of amusement, are associated with a greater

number of Duchenne smiles (Ekman et al., 1990; Frank, Ekman, & Friesen, 1993). Second, as demonstrated by Frank and colleagues (1993), the Duchenne smile is important with respect to impression formation; individuals exhibiting Duchenne smiles, as opposed to non-Duchenne smiles, are more likely to be attributed as having positive social characteristics (e.g., being more natural, outgoing, and relaxed). Although the extent to which the expression of the Duchenne smile is the result of socialization or "hard-wiring" is not totally clear, the observation that infants are more likely to exhibit a Duchenne smile when approached by their mother, rather than a stranger, provides some indication that the Duchenne smile might be a biologically based and, to some extent, universally evolved display of emotion (Dickson, Walker, &Fogel, 1997; Fox & Davidson, 1988). In addition to the behavioral and potential evolutionary significance of the Duchenne smile, Ekman et al. (1990) reported that Duchenne smiles, compared to other smiles, were associated with greater left-sided anterior temporal and parietal EEG activation. Finally, the uniqueness of the Duchenne smile is further underscored by Root and Stephens's (2003) finding that the central peak size in EMG cross-correlograms (i.e., an indicator of common synaptic drive) was larger during the co-contraction of the orbicularis oculi and the zygomaticus major during a posed Duchenne smile than during posed expressions of sadness and horror or than during three other non-naturalistic "contrived" expressions. As noted by Root and Stephens (2003), the finding that the components of posed smiles exhibit stronger common synaptic drive than do posed expressions of sadness or horror might be accounted for by the fact that people are often called upon via social display rules to voluntarily produce a smile and are less frequently required to voluntarily pose sadness or horror.

The exact relationship between smiling and subjective emotional experience is unclear. Smiling occurs in a variety of social settings (Ekman, Friesen, O'Sullivan, 1988; Fridlund, 1991;

Keltner, 1995; Kraut & Johnston, 1979; Lafrance& Hecht, 1999; Provine& Fischer, 1989; Sousignan&Schaal, 1996), and positive emotion, in and of itself, does not appear to be necessary or sufficient for the production of a smile (Fernandez-Dolls & Ruiz Belda, 1997; Fridlund, 1994). Nevertheless, overall, the relevant literature suggests that the Duchenne smile is a valuable behavioral marker.As stated by Soussignan (2002), "…although experience of pleasure or enjoyment is not necessarily linked to the expression of smiles (e.g., Fernandez-Dols& Ruiz-Belda, 1997), available data support the view that it is heuristically useful to make a distinction between Duchenne and non-Duchenne smiles, as they have distinct experiential, situational, and cerebral correlates" (p. 55).

Smiling Behavior in PD

Using both FACS and impressionistic ratings, investigators have found that people with PD exhibit fewer Duchenne smiles relative to controls (Simons et al., 2003; Smith et al., 1996). Pitcairn and colleagues (Pitcairn, Clemie, Gray, &Pentland, 1990**)** found that people judged individuals with PD, relative to controls, as displaying fewer "happy" smiles, which, based on the anatomically-based descriptions within the text of their paper, appear to correspond to Duchenne smiles, and more "false" smiles. Interestingly, the total number of smiles appears to be approximately equal for PD and controls (Pitcairn et al., 1990; Simons et al., 2003; Smith et al., 1996).

The finding that Duchenne frequency decreases in individuals with PD, whereas the total number of smiles appears to be approximately equal for PD and controls is not well understood. One possibility is that PD participants are aware of their diminished spontaneous smiling (i.e., Duchenne smiling) and attempt to compensate with voluntary (i.e., posed or non-spontaneous) facial expressions (Halfacre et al., 2009). The finding that individuals with PD and without PD

provide comparable ratings when asked to describe their own facial expressivity (Borod et al., 2008; Simons et al., 2004) is compatible with this account and lends further support to the notion that individuals with PD might recognize a deficit in spontaneous smiling, actively attempt to compensate using deliberate smiles, and might believe that their compensatory efforts have been successful (Halfacre et al., 2009).

There is, however, an alternative explanation for the observed similarity in terms of total smiling but a difference in terms of the proportion of Duchenne versus non-Duchenne smiling. It is possible that people with PD, whether or not they are aware of their deficit in spontaneous emotional expression, do not actively compensate for diminished spontaneous smiling. Instead, due to PD-related pathology, it may be that spontaneously generated smiles in PD often take on a "false"/non-spontaneous appearance due to the absence of cheek raise/eye-corner wrinkles. In more objectively defined terms, it is possible that PD differentially affects the two components of the Duchenne smile such that the zygomatic major muscle group (in FACS terminology, AU12, which is responsible for lip stretching) is relatively preserved whereas the orbicularis oculi (in FACS terminology, AU6, which is responsible for cheek raise/eye corner wrinkling) is absent or diminished.

Potential Mechanisms Underlying PD-Related Changes in Smiling

A re-examination of the communicative and neuroanatomical correlates of the smile may allow for a better understanding of the potential neural mechanisms responsible for the observed changes in smiling behavior in PD. As previously mentioned, the Duchenne smile tends to reflect subjectively felt positive emotion whereas non-Duchenne smiles are less directly tied to the subjective experience of positive emotion. Additionally, recall that a basic non-Duchenne smile has only one component (AU12 [basic smile] involvement), whereas a Duchenne smile has

two components (AU12 [basic smile] and AU6 [check raise/eye-corner wrinkling] involvement). Thus, the basic smile predominantly recruits lower face musculature, and the Duchenne smile additionally recruits upper face musculature.

Theoretically, a change in intensity or frequency of Duchenne smiles in PD can arise for a number of (not necessarily mutually exclusive) reasons. One possibility is that PD differentially impacts the neural pathways innervating the upper and lower face. Different neuroanatomical connectivity is associated with the lower and upper face (Matsumoto, Keltner, Shiota, O'Sullivan, & Frank, 2008), and it is possible that the progression of PD exerts a disparate impact upon these two facial regions. Another possibility is that the decreased number of Duchenne smiles in PD is an artifact of relatively greater impairment of spontaneous emotional expression as opposed to posed emotional expression. If the typical Duchenne smile is indeed generated via spontaneous activation and the typical non-Duchenne smile is generated via more deliberate control, then we should expect to see a greater reduction in Duchenne as opposed to non-Duchenne smiling. Finally, another possibility, is that, in PD, reductions in Duchenne smiling, as well as diminished intensity/frequency of other emotional expressions, might in part reflect altered or diminished positive emotional experience.

Clinical Considerations: Implications and Treatment of Facial Expression Deficits in PD

Clinical and Psychosocial Significance

In social contexts, facial expressions serve at least two broad functions. On the one hand, the "receiver" or "observer" of non-verbal communication uses this information to infer the "sender's" cognitive/emotional state. In addition to being an indicator of subjective experience, facial expressions play a causal role in the social dynamic; our impressions of another person

(e.g., our estimations of their intellect and personality) and, consequently, our behavior toward that individual, can be shaped by their non-verbal communication.

Pitcairn and colleagues (1990) note that in PD and certain other neurological conditions, "...the outward behavior may change, but the internal mood or state does not alter, leading to a mismatch between the feelings of the actor and the impression they give to others" (p. 177). In other words, patients with masked facies and other deficits of non-verbal expression are at risk for miscommunication, that is, they may send messages that do not accurately reflect their cognitive and emotional state. The result of this "mismatch" can be both confusing and troubling with respect to interpersonal communication (Brozgold et al., 1998; Katsikitis&Pilowsky, 1996; Pentland, Pitcairn, Gray, & Riddle, 1987; Pitcairn et al., 1990).

Notably, novice and even experienced clinicians have a tendency to develop—at least initially—inaccurate and unfavorable impressions of individuals with PD (Monrad-Krohn, 1957; Pentland et al., 1987, 1988; Tickle-Degnen& Lyons, 2004). For example, Pentland and colleagues (1987) had both students and experienced healthcare workers view silent video footage of PD patients and controls (cardiac ischemia patients) and found that both groups rated PD patients, relative to controls, as less intelligent and more anxious, hostile, suspicious, unhappy, bored, and tense. Summarizing their findings, Pitcairn and colleagues (1990) note that PD patients, quite simply, were generally seen as "much less likeable" (p. 178). Underscoring the likelihood that these impressions did not accurately reflect the inner qualities of PD patients in the study by Pentland et al. (1987), there were no differences between the PD patients and controls on measures of mood/anxiety (Spielberger's State-Trait Anxiety Inventory; Beck Depression Inventory), personality (Cattell's 16 Personality Factor questionnaire; Hostility Scale

of the Multiple Affect Adjective Checklist), and/or intelligence (Cattell's primary factor B [a measure of verbal reasoning ability]).

Finally, although much of the above review was discussed in terms of discrete emotions, distinct modes of processing, and separate channels of emotional expression, it should be underscored that, in the context of an individual's day-to-day functioning, these components of emotional functioning are highly interconnected. Accordingly, even "isolated" deficits in emotional processing might have considerable direct and indirect impact on an individual's ability to engage in emotionally meaningful interpersonal experiences and to produce socially/emotionally appropriate facial behavior. For example, an isolated deficit in emotion recognition might lead to difficulty identifying and responding to the emotional needs of others. Furthermore, reduced ability to recognize emotion in one's environment might, in essence, reduce the frequency and/or intensity of emotional "input," thus decreasing the probability and/or intensity of subjective emotion. Finally, an isolated emotion recognition deficit might result in an overall "disengagement" from external emotional stimuli such that subjective emotional experience is driven more by internally generated emotional stimuli (e.g., memories). In short, emotional functioning, like many other psychological phenomena, is a complex, dynamic, process that emerges from interconnected subprocesses.

Treatment

Facial expressivity, along with other PD symptomatology, has been observed to improve with a number of interventions including levodopa treatment (Treciokas, Ansel, & Markham, 1971), deep brain stimulation (DBS; Chung, Jeon, Kim, Sung, & Lee, 2006), nicotine administration (Fagerstrom, Pomerleau, Giordani, &Stelson, 1994), electroconvulsive therapy (ECT; Lebensohn& Jenkins, 1975), picoTesla range magnetic fields (Sandyk&Derpapas, 1993),

and behavior-based interventions (e.g., Katsikitis& Pilowsky,1996). The focus of this review is a behavior-based intervention, the Lee Silverman Voice Treatment (LSVT LOUD; Ramig, Pawlas, et al., 1995: Ramig et al., 2001).

LSVT LOUD is one of the few behavioral interventions that, based on anecdotal and preliminary empirical evidence (Spielman, et al., 2003), appears to have a positive effect on facial expression in PD. In their investigation, Spielman et al. (2003) had viewers rate video footage (collected as part of a study on speech/voice treatment efficacy in PD) of 44 individuals with PD in terms of their facial muscle activity as well as their efficacy of communication (i.e., the degree to which the rater perceived the individual as "engaging" in terms of interpersonal communication). PD participants received either LSVT LOUD, a high effort voice treatment directed at the phonatory system, or an alternative treatment, matched on all variables but directed to the respiratory system; video footage was obtained at baseline and after treatment. Compared to the PD group receiving respiratory treatment, the PD group receiving LSVT LOUD generally received more ratings of both increased facial mobility and increased "engagement." Interestingly, LSVT LOUD was developed to target voice and speech deficits, and any benefit to facial expressivity was an initially unexpected collateral gain.

To better understand the mechanisms by which LSVT LOUD might affect facial expression, it is necessary to briefly review both the speech/voice deficits present in PD, as well as the theory and practice of LSVT LOUD treatment. About 75-89% of people with idiopathic PD have voice and speech disorders (for review, see Fox et al., 2002; for discussion of prosody deficits in PD and other conditions, see McCabe et al., 2011). Voice and speech dysfunction in PD are characterized by hypophonia, diminished prosody, breathy/hoarse voice, and articulatory imprecision. Whereas a variety of medical interventions (e.g., neuropharmacological and

neurosurgical) have demonstrated efficacy in reducing the severity of motor symptoms (e.g., tremor, bradykinesia, and rigidity), the efficacy of these treatments on voice and speech production has been less consistent (for review, see Fox et al., 2002; Schulz & Grant, 2000). Communication difficulties arising from speech/vocal changes in PD can be further compounded by the presence of masked facies; considering the fact that PD can compromise two distinct communication channels (speech and facial affect), it is not surprising that reduced ability to communicate is reported to be a major obstacle by both individuals with PD and their family members (Fox et al., 2002).

LSVT LOUD is a behavior-based speech-language intervention targeting vocal loudness. Conducted over the course of four individual sessions per week for four weeks, LSVT LOUD therapy involves "high effort" speech production, repetition, enhanced sensory awareness, intensive treatment delivery, and quantification of behaviors (Ramig, Pawlas, et al., 1995). With respect to the "sensory awareness" component of LSVT LOUD, the individual with PD is assisted in monitoring and calibrating his or her speech volume such that normal loudness is achieved. The necessity of this training component stems from the fact that individuals with PD often feel that they are shouting when speaking at volumes that others perceive to be normal. The efficacy of LSVT LOUD on speech/vocal outcomes has been documented in a number of studies (e.g., Baumgartner, Sapir, & Ramig, 2001; Ramig, Countryman, O'Brien, Hoehn, & Thompson, 1996; Ramig, Countryman, Thompson, & Horii, 1995; Ramig et al., 2001; Sapir, Spielman, Ramig, Story, & Fox, 2007; Narayana et al., 2009; Liotti et al., 2003).

As noted above, despite having been developed for the treatment of speech/voice deficits, available evidence suggests that LSVT LOUD also affects facial expressivity (Spielman et al., 2003). Although the mechanism(s) by which LSVT LOUD might impact facial expression is

unknown, Spielman et al. (2003) discuss a number of possibilities. One explanation posits that vocalization and facial expression, despite being distinct modalities at the level of output, are actually neurologically yoked as a result of synchronous activation during emotional expression (e.g., Banse& Scherer, 1996). To the extent that face and voice are "neurologically yoked," it would be reasonable to expect that the training of one pathway (i.e., voice) would confer an indirect training effect on the second pathway (i.e., facial expression). The precise mechanism by which this indirect effect might be accomplished is unknown, but the available evidence implicates peripheral neural substrates, such as the facial nerve, or more central neural structures, such as the anterior cingulate (Spielman et al., 2003).

Neural coupling of some kind between voice and face may indeed account for post-LSVT LOUD improvements in facial expression, but psychological as well as motoric explanations should also be considered. Concerning a potential psychological mechanism to account for improved facial expression following LSVT LOUD, it is possible that a patient's satisfaction with vocal treatment is reflected through his or her facial affect (i.e., the patient is pleased with vocal treatment and, consequently, appears visibly happier/more socially engaged). With respect to a purely motoric explanation, post-treatment strengthening of orofacial musculature, which is documented with LSVT LOUD (for discussion, see Fox et al., 2002), might also account for increased facial expressivity.

This notion that increased muscle strength may result in improved facial expression is reflected in other treatments targeting masked facies. Katsikitis and Pilowsky (1996), for example, investigated an "orofacial physiotherapeutic treatment" (OPT). Sixteen individuals with PD were allocated to either the OPT treatment condition or designated as controls. Facial assessments (based on random frames of video-recordings, which were obtained during a brief

interview) were obtained at baseline, post-treatment, and follow-up (four weeks post-treatment). Individuals assigned to the treatment condition were found to exhibit and maintain improvement in mouth-opening, lip mobility, and eyelid activity. Facial mobility changes were not assessed in emotional/affective terms but the authors note that improved facial mobility may allow for improved facial affect in PD.

Additionally, ongoing research includes a Phase II Clinical Trial being conducted at the University of Florida (Bowers, Sapienza, Okun, & Fernandez), which is investigating the effect of "high intensity respiratory muscle strength training" (http://clinicaltrials.gov/ct2/show/NCT00350402; study ID number R01NS050633) on facial expression deficits in individuals with PD. The treatment is described as behavioral and will utilize a device (a mouthpiece engineered to provide respiratory resistance) intended to strengthen the muscles used in facial expression. The study will be employing both objective measures (computerized measures of facial movement) and subjective measures to evaluate, respectively, the impact of treatment on facial expressivity, emotional variables, and quality of life.

Proposed Research

Although developed as a voice and speech intervention for individuals with Parkinson's disease, preliminary evidence suggests that the Lee Silverman Voice Treatment (LSVT LOUD) might also reduce the facial expressivity disturbances that can occur in this population. Long-term follow-up data, however, has yet to be obtained for the efficacy of LSVT LOUD in treating facial expressivity deficits. Furthermore, little is known about the specific changes in facial expressivity that might result from treatment.

Using video footage of healthy controls and individuals with PD, The Emotion Lab at Queens College of the City University of New York, is currently investigating the effect of two behavior-based (i.e., non-surgical, non-pharmacological, and non-psychotherapeutic[1]) voice treatments on facial expressivity in PD. One of these treatments, Artic, targets vocal articulation. The other treatment, the Lee Silverman Voice Treatment (LSVT LOUD), targets vocal loudness. The current study complements the broader project through an examination of the nature, magnitude, and duration of facial expressivity changes following treatment. Specifically, using the Facial Action Coding System (FACS; Ekman, Friesen, & Hager, 2002), a method that, to the best of our knowledge, has yet to be employed in the evaluation of facial expressivity interventions for PD, this study addressestwo aims: 1.) evaluation of the long-term (6-month) efficacy of LSVT LOUD compared to an articulation-based intervention,[2] and 2.) evaluation of differences in smiling behavior as a function of treatment.

Materials and Methods

Participants

The individuals who appear in the video footage are healthy controls (age and sex matched; n = 11) and individuals with Parkinson's disease (n = approximately 39). The PD group comprises three treatment conditions: individuals receiving no treatment (n = approximately 13), individuals receiving an articulation-based treatment (Artic; n = approximately 12), and individuals receiving the Lee Silverman Voice Treatment (LSVT LOUD; n = approximately 14). Group demographic characteristics do not significantly differ in terms of

[1] While these interventions might be psychotherapeutic to some degree (e.g., insofar as the outcome of treatment improves quality of life), they target discrete components of speech production.

[2] The immediate post-treatment effect of LSVT is already currently under investigation within the Emotion Lab at Queens College.

age, sex, ethnicity, or education (in years), and PD groups do not differ in terms of time-since-diagnosis or Hoehn-Yahr stage (Hoehn &Yahr, 1997), as demonstrated by ANOVAs and chi-square tests with *p*-values exceeding .05.

Inclusion/Exclusion Criteria

Participants were excluded from the study if they manifested moderate to severe dementia or severe depression. The cutoffs for these determinations were:

- Moderate to severe cognitive impairment: indicated by scores < 25 on the Mini Mental State Examination (MMSE; Folstein et al., 1975).
- Severe depression: indicated by scores >24 on the Beck Depression Inventory, 2nd version (BDI-II; Beck, 1967).

Video Recordings

The video recordings used in this study were produced in the context of the larger study, described below. These recordings show participants engaged in monologues intended to elicit a feeling of happiness. To accomplish this, participants were asked to recount a recent happy event in their lives. Each participant engaged in the happy monologue procedure on three different occasions: Time 1(baseline), Time 2 (post-treatment), and Time 3 (6-month follow-up).

The procedure for eliciting the monologues was developed by Dr. Joan Borod (Queens College, City University of New York) and is adapted from the New York Emotion Battery (Borod et al., 1992). The video recordings of these monologues were produced in the lab of Dr. Ramig as part of an NIH-funded study evaluating the clinical efficacy of the LSVT LOUD (grant R01-DC01150).

Although the NIH-funded protocol included neutral monologues and three emotional monologues (Happy, Sad, and Angry) of varying length (almost all were between one and three

minutes), the labor intensiveness of FACS coding—used in the currentstudy—requires that coding be restricted to a subset of the original video recordings. Previous research in the Queens College Emotion Lab investigated pre-/post-treatment findings for Happy monologues. The current research extends these findings by providing follow-up data for the Happy monologues. Additionally, this study, for select facial expressions, provides data on expressive intensity and temporal aspects of expression (described in detail in the Monologue Coding section, below). In order to acquire follow-up data that is methodologically comparable to previously coded baseline and post-treatment data, the video selection protocol was identical to that used by Dumer and colleagues (Dumer et al., 2010). Specifically, coded footage includes the final 60 seconds of Happy monologues.[3]

Monologue coding

Three types of coding data (see subsections below for details) were extracted from the monologue footage. All video footage was coded for the occurrence of *Action Units* (AUs; Ekman, Friesen, & Hager, 2002), the basic units of facial expression described in FACS. Additionally, intensity coding and temporal coding was conducted for two AUs of interest, AU6 ("Cheek Raiser and Lid Compressor") and AU12 ("Lip Corner Puller"). These AUs were selected due to their role in two categories of smiles, Duchenne smiles (AU6 co-occurring with AU12) and non-Duchenne smiles (AU12 without AU6). The psychosocial significance of these categories of smiles is described in the literature review section of this proposal. Anvil, a video

[3] The decision to use the final 60 seconds of monologues was made in light of the finding by Kazandjian, Borod, and Brickman (2007) that individuals participating in the NYEB battery are generally more expressive in the middle third and latter portion (i.e., last third) of monologues. The decision to use Happy monologues was made because the investigators believed that Happy monologues would potentially yield more facial expressivity and thus more coding data. Happy monologues might also provide the most psychosocially significant data set, given the interpersonal value of being able to appropriately express positive emotion (for review of smiling behavior, see LaFrance, Hecht, &Paluck, 2003).

annotation program (developed by Michael Kipp; Kipp, 2001), was used to view the digitized monologues and to record coding data.

Coding for AU occurrence and intensity. The Facial Action Coding System (FACS; Ekman, Friesen, & Hager, 2002) was used to code for the occurrence of facial expressions (*AU occurrence*) as well as the intensity of facial expressions (*AU intensity*). In FACS, the occurrence of any facial expression can be represented as a combination of visually observable, objectively defined, facial movements termed *Action Units* (AUs). AU intensity can be coded using a 5-point scale wherein "A" is used to reflect barely discernible AU involvement, "E" is used to reflect maximal AU activation, and "B" through "D" are used to reflect intermediate degrees of activation.Behavioral markers to guide intensity coding are described in the FACS manual). FACS coding criteria and methods are described in detail in the FACS training manual (Ekman et al., 2002). A FACS-certified coder (A.D.) has coded AU occurrence at Time 1 and Time 2. Another FACS-certified coder (D.M.) coded AU occurrence at Time 3 as well as AU intensity, for AU6 and AU12, at all time points.

Coding of AU occurrence at Time 1 (pre-treatment) and Time 2 (post-treatment) has, in the course of previous research in the Queens College Emotion Lab, already been completed by FACS coder A.D. (Dumer et al., 2010). When coding AU occurrence at Time 3 (6-month follow-up), the order of monologue coding was quasi-randomized such that the first and second halves of the coding schedule were balanced in terms of participant group (LSVT LOUD, Artic, No Treatment, and Healthy Control). When coding AU intensity, the first and second halves of the coding schedule were balanced in terms of participant group (LSVT LOUD, Artic, No Treatment, and Healthy Control) and time (Time 1, Time 2, and Time 3).

Temporal coding. A FACS-certified researcher (D.M.) coded temporal features of AUs

6 and 12 at Times 1 (baseline) and 2 (post-treatment).[4]The process used to record AU *occurrence* requires the coder to record an AU's onset (the moment an AU first appears) and offset (the final visible moment of an AU). To evaluate *temporal characteristics* of AU activation, two additional behavioral time points were recorded: the *first peak* of an AU (or the moment when an AU first reaches a plateau) and its *highest peak* (i.e., *apex*, which, for some AUs, is identical to the first peak). By recording temporal data, one can assess hypotheses pertaining to bradykinesia at the stage of motor execution.[5] Furthermore, coding these two time points provides basic information about the contour of an AU's onset phase (for example, one could assess whether the onset phase of an AU was comprised of a single relatively smooth rise in intensity or if the AU proceeded to peak intensity via a step-wise increase in intensity).

For the temporal coding of AUs 6 and 12 at Time 1 and Time 2 (for which basic FACS coding was completed by Dumer et al., 2010), the first and second halves of the coding workload were balanced in terms of participant group (LSVT LOUD, Artic, No Treatment, and Healthy Control) and time (Time 1 [baseline] and Time 2 [post-treatment]).

As per the planned coding procedure, facial movements judged to be unrelated to emotional expression (i.e., artifacts of speech articulation or other non-affective actions such as wetting one's lips, tics, etc.) were not coded. Prior to calculating inter-rater reliability, D.M. identified coding disagreements that appeared to be due to differences in judging which facial actions were emotional (and thus coded) versus non-emotional (and thus not coded).

---

[4]Time 3 will not be subject to temporal analysis. This decision is due to the fact that FACS coding at Times 1 and 2 was collected by one coder (A.D.) whereas FACS coding at Time 3 (follow-up) was completed by a different coder (D.M.). Because even subtle inter-coder differences in onset coding could obscure temporal data, limiting temporal data to Time 1 and Time 2 is thought to be a more conservative approach to investigating possible temporal changes.

[5] Notably, in contrast to other studies, this measure of bradykinesia will not reflect reaction time delays between stimulus and AU onset.

Disagreements, involving facial actions which could be reasonably interpreted as either non-emotional or marginally emotional, were resolved by adjusting DM's coding such that it was calibrated to AD's coding; disagreements were resolved in this way to ensure that the final reliability calculations reflected only those differences in *detecting* emotional facial expressions (i.e., excluding disagreement related to *judgments* as to the emotional versus non-emotional nature of a facial action).

Inter-Coder Reliability

Inter-coder reliability was separately calculated for AU occurrence, AU intensity, and temporal coding. Descriptions of the inter-coder reliability procedures are provided below.

Reliability for AU occurrence coding. Prior to coding the video footage to be used in the study, two FACS-certified coders (D.M. and A.D.) coded a subset of monologues in order to assess inter-coder differences. Coding disagreements within this preliminary dataset were discussed to minimize the possibility that coder idiosyncrasies could contaminate the final dataset.

Coding of AU occurrence at Time 1 (pre-treatment) and Time 2 (post-treatment) has already been completed by FACS coder A.D. (Dumer et al., 2010). To establish inter-coder reliability for these monologues, D.M. coded a subset (24 of 112 monologues [approximately 20%]) of the files coded by A.D. In the proposed study, AU occurrence at Time 3 (follow-up) was coded by D.M. To establish inter-coder reliability for AU occurrence in these monologues, A.D. coded approximately 20% of Time 3.

For AU occurrence, three types of inter-coder reliability values were calculated: one for overall agreement, one for Duchenne smiles, and one for non-Duchenne smiles.Inter-coder reliability for the coding of AU occurrence was calculated using the formula described in the

Investigator's Guide to the Facial Action Coding System (Ekman, Friesen, & Hager, 2002).[6]

Reliability for AU intensity coding. Intensity data were coded by D.M. A subset (20%) of the intensity data were coded by A.D. in order to establish inter-coder agreement. A modified version of the formula described in the FACS Investigator's Guide to the Facial Action Coding System was used to calculate reliability for intensity.[7]

Reliability for temporal coding. Temporal reliability was established in two stages. First, a research assistant was trained to code the temporal characteristics of AUs and coded a set of monologues that were not entered into the final analyses. D.M. qualitatively reviewedthe research assistant's temporal coding of these monologues and calculated inter-coder reliability for the training data set. Upon achieving satisfactory reliability, the research assistant proceeded with coding 20% of the inter-coder reliability data set. Temporal reliability was calculated as the number of agreements divided by the sum of the number of agreements and disagreements.[8]

Description of Dependent Variables

Aim 1dependent variables. To evaluate the long-term (6-month post-treatment)

---

[6] For a given AU configuration, inter-coder reliability = (# of AUs upon which both coders agree) x 2 / (total # of AUs scored by both coders).

[7] Intensity agreement index = (# of Agreements)/(# of Agreements + # of Disagreements), where agreement is defined an instance of intensity coding in which D.M.'s recorded intensity is equal to or +/- 1 level of intensity relative to A.D.'s coding. For example, if D.M. were to label an AU as B-level intensity and A.D. were to label an AU as exhibiting C-level intensity, this would count as an agreement. If however, D.M were to label an AU as exhibiting B-level intensity, and A.D. were to label this AU as exhibiting D-level intensity, this would count as a disagreement.

[8] Temporal agreement index = (# of Agreements)/(# of Agreements + # of Disagreements) where agreement is defined as an instance of temporal coding in which both coders determine that a time point (first peak or highest peak) occurred within a specified window of time (i.e., tolerance window). For example, with a tolerance window of 0.25 seconds, if D.M. were to code First Peak as occurring at 55.00 seconds into a monologue and M.S. were to code First Peak as occurring at 55.24 seconds into a monologue, this would count as agreement. Temporal agreement for a single data set can be calculated separately for multiple tolerance windows (e.g. 0.5 seconds, and, a more stringent criterion, 0.25 seconds).

efficacy[9] of behaviorally-based speech interventions (Artic and LSVT LOUD) in treating facial expressivity deficits in PD, several dependent variables—AU Complexity, AU Lability, AU Variability, and AU Time—were obtained for each monologue. Definitions of these variables are as follows:

AU Complexity—The average number of AUs in an AU configuration, excluding frames in which there are no AUs (e.g., if a monologue contained the following three AU configurations–AU5, AU6+AU7, and AU6+AU7+AU8–then AU Complexity would equal 2).

AU Lability—The total number of facial events (e.g., if one configuration is initially present and is followed by 5 other configurations, then AU Lability would equal 6).

AU Variability—The number of unique AU configurations (e.g., if a monologue contained the following sequence–AU1+AU2, AU1+AU3, AU1+AU2, AU1+AU3–then AU Variability would equal 2).

Aim 2 dependent variables. To evaluate differences in smiling behavior as a function of Treatment and Group(healthy control or PD), the following variables were obtained for each monologue:

Number of Duchenne Smiles—The number of co-occurrences of AU6 and AU12.

Number of Non-Duchenne Smiles—The number of times that AU12 is present in the absence of AU6.

Duchenne Index (DI)—The number of Duchenne Smiles divided by the sum of Duchenne and non-Duchenne smiles (i.e., Duchenne/(Duchenne + non-Duchenne)). The number of Duchenne Smiles minus the number of Non-Duchenne smiles, divided by the sum of Duchenne

[9] The immediate post-treatment effect of LSVT on facial expression is already currently under investigation within the Emotion Lab at Queens College.

and non-Duchenne smiles (DI = [Duchenne Number + Non-Duchenne Number]/[Duchenne Number + Non-Duchenne Number]).

Smile Intensity (Duchenne)—The average intensity of AU 12, when co-occurring with AU6.[10]

Smile Intensity (Non-Duchenne)—The average intensity of non-Duchenne smiles within a monologue. Only AU12 intensity was coded (by definition, AU6 was absent; other co-occurring AUs were not be scored).

Smile Intensity (Overall)—The average intensity of smiles, AU12, regardless of whether the Duchenne marker, AU 6, is present.

Time to Apex (Duchenne)—The average time it takes AU12 to reach peak intensity, when AU 12 co-occurs with AU6.

Time to Apex (Non-Duchenne)—The average time it takes AU12, in the absence of AU6, to reach peak intensity.

Time to Apex (Overall)—The average time it takes a smile to reach peak intensity, averaged across Duchenne and non-Duchenne smiles.[11]

Hypotheses

Aim 1—Assessing Treatment Efficacy at Follow-Up (Time 3) UsingGlobal Measures of Facial Expressivity

[10] On an exploratory basis, AU6 intensity was also recorded. AU6s were included in the intensity scoring regardless of their co-occurrence or independence of AU12.

[11] *Time to Apex* was the primary mode of describing temporal features of smiling execution. However, as disease status or treatment may affect the slope of AU12 onset, the proposed research will note if the rising phase (i.e., onset to apex) of AU12 is comprised of a single peak (the apex) or multiple peaks (momentary peaks and/or plateaus preceding the apex). If multiple peaks are identified within the rising phase, the time of the first peak (in addition to the apex) was recorded to allow for exploratory analyses.

Aim 1 assessed the long-term efficacy of behavior-based speech treatments on facial expressivity in PD. Itwas hypothesized that at Time 3 (i.e., 6-month follow-up), both ARTIC and LSVT LOUD wouldresult in statistically significant increases in facial expressivity compared to Time 1 (pre-treatment). Furthermore, it was predicted that the effect of LSVT LOUD would be significantly greater than the effect of Artic. Finally, because there was no active treatment between Time 2 (post-treatment) and Time 3 (follow-up), it was predicted that treatment effects (i.e., increased facial expressivity) of both ARTIC and LSVT LOUD woulddiminish such that the values of the Aim 1 outcome measures at Time 3 would be significantly less than the values obtained at Time 2

Aim 2—Smiling Behavior in Treated and Untreated Parkinson's Disease

Aim 2 evaluated how PD and speech treatment affect the number, intensity, and temporal features of Duchenne and non-Duchenne smiles.

Number of smiles at baseline. In accord with previous studies, it wasexpected that, relative to controls, individuals with PD wouldexhibit a lesser amount of Duchenne smiling. This expectation was evaluated by testing two hypotheses. At Time 1 (pre-treatment), it washypothesized that, on average,individuals with PD (comprised of all PD participants at Time 1) wouldproduce fewer Duchenne smiles, as compared to healthy controls. Second, at Time 1 (pre-treatment), it waspredicted that the Duchenne Index wouldbe smaller for the PD group compared to Healthy Controls (i.e., the ratio of Duchenne to non-Duchenne smiles wouldbe smaller in the pre-treatment PD participants, relative to Healthy Controls). In light of the previous findings, a one-tailed test was used to evaluate baseline differences in Duchenne Number and Duchenne Index. All other baseline frequency hypotheses were evaluated using

two-tailed tests.

Number of smiles after treatment. It waspredicted that LSVT LOUD, and, to a lesser extent, Artic, wouldresult in an increased number of smiles. Furthermore, it waspredicted that LSVT wouldincrease the Duchenne Index (i.e., increase the number of Duchenne smiles relative to non-Duchenne smiles) whereas ARTIC wouldnot. If this differential effect of treatment on Duchenne versus non-Duchenne smiling were to beobserved, it would provide support for the hypothesis that the effect of LSVT LOUD is not restricted to speech-specific AUs, but instead affects both speech *and* non-speech AUs (e.g., AU6, cheek-raiser, an upper-face AU) via activation of motor pathways shared by both speech and non-speech output (see Fox et al., 2006 for discussion of potential mechanisms of action). Outcome measures to assess this hypothesis include the Duchenne Index and the Number of Smiles (i.e., sum of both Duchenne and non-Duchenne smiles).

Intensity of smiles at baseline. It wasexpected that, relative to controls, individuals with PD would exhibit less intense smiles (Smith et al., 1996)). This hypothesis was tested using one-tailed tests to compare average smile intensity PD participants to healthy controls at Time 1.

Intensity of smiles after treatment. It waspredicted that LSVT LOUD, and to a lesser extent, Artic, wouldresult in increased smile intensity from Time 1 (pre-treatment) to Time 2 (post-treatment), and that any treatment effects observed at Time 2 would be diminished at Time 3 (follow-up). As described above, it was hypothesized that the effect of LSVT LOUD wouldnot be restricted to speech-specific and adjacent AUs, but would instead affect both speech *and* non-speech AUs. This hypothesis would be supported if LSVT LOUD were to result in increased intensity of AU6 (cheek raiser; a marker of emotional expression that is not required for speech) at Time 2, compared to Time 1.

Supplementary Analyses

Temporal features of smiling in PD and healthy controls. In accordance with previous literature (Bowers, Miller, Bosch, et al., 2006), it waspredicted that PD participants wouldddemonstrate bradykinetic facial expression. This hypothesis would be supported if Time to Apex were to besignificantly greater in the PD group, compared to healthy controls.

The impact of treatment on temporal features of facial expression. The potential impact of treatment on the duration of facial expression onset (i.e., Time to Apex) is unclear. All else being equal, if treatment increasesexpressive Intensity, one would expect correspondingly increased Time to Apex, assuming a constant onset speed (i.e., assuming a constant speed of muscle contraction). On the other hand, treatment might result in more efficient (i.e., more rapidly executed) facial expression.

Results

Group Demographics

Group demographics are summarized in Table 1, Table 2, and Table 3.

Healthy controls versus individuals with PD.As expected, baseline BDI scores significantly differed between Healthy Controls (HC; *Mdn* = 1.00, *M* = 2.27) and individuals with PD (*Mdn* = 8, *M* = 9.04), *U* = 431.00, *z* = -3.79, *p* = <.001, as revealed by a Mann Whitney test. Groups did not significantly differ in Age, *U* = 159.00, *z* = -1.83, *p* = .07. The PD group was significantly older than the Healthy Control group, $F(1,54) = 4.11$, $p = .05$. There was no significant difference in years of education between the two groups, *U* = 193.00, *z* = -1.15, *p* = .25. There was a significant difference in MMSE scores, *U* = 120.50, z = -2.76, *p* = .01, such that Healthy Controls scored higher (*M* = 29.73) than individuals with PD (28.73).

Equality of Treatment Groups.As expected, BDI score differered across treatment groups (HC, PD-Untreated, ARTIC, and LSVT),as demonstrated by a Kruskal-Wallis analysis of variance by ranks, $\chi^2(3)$, $p = <.01$ (See Tables 10 and 11 for correlations between BDI and dependent variables). Pairwise comparisons revealed that the Healthy Control group had lower BDI scores, as compared to each of the PD treatment groups; there were nosignificant differences between the PD groups. Age did not significantly differ between treatment groups, $F(1,52) = 1.84$, $p = .15$. Years of education did not differ among treatment groups, $F(1,52) = 5.62$, $p = .61$.

MMSE score did significantly differ, marginally, between treatment groups, $\chi^2(3)$, $p = .055$, such that Healthy Controls scored slightly higher on average ($M = 29.73$) as compared to the PD groups (Untreated PD = 28.88); ARTIC = 28.75, LSVT = 28.56).

Regarding PD-specific variables, the PD groups did not significantly differ in terms of Hoehn &Yahr stage, $\chi^2(2) = .52$, $p = .77$, or Time Since Diagnosis, $\chi^2(2) = .71$, $p = .70$.

Reliability

Reliability data are summarized in Table 4 and Table 5. Additional details are described below.

Aim 1. Global Measures of Facial Expressivity. Inter-coder reliability of global measures of facial expressivity was examined in two stages: first, for the coding of baseline (Time 1) and post-treatment data; and second, for the coding of six-month follow-up data (Time 3). Reliability measures included overall coding agreement, assessed by percentage of event-based agreement and FACS Agreement index (described above), as well as intra-class correlation coefficients (ICCs) to determine variable-specific agreement for each of the three Aim 1 outcome measures (Lability, Variability, and Complexity).

Times 1 and 2. As reported in Dumer (2011), for Times 1 and 2, intra-class correlation coefficients for individual Aim 1 outcome measures were .80 for AU Lability, .75 for Variability, and .60 for AU Complexity. Percent agreement on the occurrence of an event was 81% and the FACS Agreement Index was 0.71. For times 1 and 2, basic FACS coding (i.e., Action Unit / Event coding) was conducted by coder A.D. with D.M. coding the reliability subset (Dumer, 2011).

Time 3. For Time 3, intra-class correlation coefficients, using a consistency definition and a two-way mixed-effects model, for each of the three Aim 1 dependent variables were as follows: Lability (single-measures = .81; average measures = .90); Variability (single measures = .86; average measures = .93); and Complexity (single measures = .44; average measures = .61). The intra-class correlation coefficient procedure was repeated using an absolute agreement definition, as opposed to a consistency definition. Single measure and average measure coefficients were unchanged for Complexity; .81 and .89 for Lability; and 0.74 and .85 for Variability. Although each reliability subset for all three outcome measures met the assumption of normality as assessed by the Shapiro Wilk test (Shapiro & Wilk, 1965) (>.05)[12], outliers were present, as assessed by boxplot inspection. Specifically, there were no outliers for Lability anda single outlier in A.D.'s Variability, which corresponded to a near-outlier in D.M.'s Variability; there were 2 outliers in A.D.'s Complexity. Lability and Variability exhibited positive skew; for Complexity, positive skew was exhibited by one coder whereas the other exhibited negative skew. Transformations were performed. All data remained normal after square root transformation; there were no outliers for Lability or Variability, and there was a single outlier for A.D.'s Complexity. After log10 transformation, A.D. Complexity was no longer normal (p =

[12] In light of the sample size (i.e., <50), the Shapiro-Wilk test was used for assessing normality (Maxwell & Delany, 1990).

.04); single outliers were present in A.D.'s Complexity and Variability. A negative skew correcting square root transformation was conducted on Complexity; both A.D.'s and D.M.'s distributions were normal but a single outlier was present in the A.D. distribution. The same pattern held for a negative skew correcting log10 transformation. After a negative skew correcting inverse transformation, A.D.'s Complexity was no longer normally distributed.

Because no single arithmetic transformation would result in normal distributions and the absence of outliers, the data were rank transformed and ICCs were recalculated. After rank-transformation, ICCs, employing a consistency definition, were as follows: Lability = .80 (single-measures), .89 (average-measures); Variability = .81 (single-measures), .89 (average-measures); Complexity = .63 (single-measures), .77 (average-measures). In light of the coding methodology and distribution of the reliability coding subset, single-measures, consistency definition ICCs are thought to be the most appropriate measures of reliability; these values are reported in Table 4. Absolute agreement ICCs were not calculated for rank-transformed data.

As the reliability of Complexity coding was relatively lower, compared to the other two dependent variables, the Complexity reliability subset was inspected for possible sources of inter-coder disagreement; based on examination of the raw tabular data and a scatter plot, disagreement seemed to result from overall disagreement as opposed to a limited number of outliers or other readily identifiable coding discrepancy. Although intra-class correlations were high for Lability and Variability, inspection of descriptive statistics revealed that D.M.'s FACS coding tended to result in higher scores for each of the three Aim1 outcome measures (Lability [M = 9.50, SD=6.70], Variability [M = 5.70, SD = 3.27], and Complexity [M = 1.48, SD=.33]) as compared to AD's coding (Lability [M= 7.90, SD= 6.81], Variability [M = 3.80, SD = 3.12], and Complexity [M = 1.29, SD= 0.63]).

Overall percent agreement on the occurrence of an event was .66, and the average FACS Agreement Index was .40. Time 3 study data were coded by D.M., and A.D. coded the reliability subset. Time 3 reliability was calculated based on 10 participants, who produced a total of 106 expressive events.

As the Time 3 FACS Agreement Index reliability was lower than expected, reliability data were examined on a subject-by-subject basis to explore potential sources of inter-rater disagreement. Ranges and the number of subjects with the specified range were as follows: .00 through .25 (n = 4); > .25 through .50 (n = 3); > .50 through .75 (n = 3); and > .75 through 1.00 (n = 0). For 2 of the 4 subjects yielding FACS Agreement Indices between .00 and .25, most disagreements appeared to result from discrepant AU sensitivity, e.g., one raterwould code a single, unbroken AU12 throughout a monologue, whereas another coder would code parts of the monologue as neutral and parts of the monologue as AU12, thus resulting in multiple disagreements and, ultimately, lower inter-rater reliability.

Aim 2. Smile Behavior.

Reliability of smile-related FACS coding. Reliability of smile coding was evaluated separately for Times 1 and 2 (primary coder was A.D.; reliability coder was D.M.) and for Time 3 (primary coder was D.M.; reliability coder was A.D.); reliability was calculated for Duchenne smiles, Non-Duchenne smiles, and overall smile agreement (i.e., in which agreement was not dependent upon smile classification).

For Time 1 and 2, percent agreement, defined by inter-coder onset agreement not exceeding one second, reliability, calculated across the reliability sample, was .74 for Duchenne smiles and .46 for non-Duchenne smiles; overall smile percent agreement was .57. A second class of agreement—in which agreement was defined as any degree of same-category overlap

(e.g., if the primary rater coded a Duchenne smile that overlapped with the reliability coder's Duchenne smile, this would count as agreement)—was also calculated. Using this definition, reliability was .74 for Duchenne smiles, .73 for Non-Duchenne smiles, and .79 for overall agreement, as defined by any degree of overlap between either type of smile.

For Time 3, percent agreement, defined by inter-coder onset agreement not exceeding one second, reliability, calculated across the reliability sample, was .67 for Duchenne smiles and .32 for non-Duchenne smiles; overall smile percent agreement was .44. Agreement defined by any degree of same-category overlap was also calculated. Using this definition, Duchenne agreement was .75 and non-Duchenne agreement was .66; overall percent agreement was .74. Descriptive statistics for the smile frequency variables were visually inspected. For coder A.D., means and *SD*s were as follows: Total Smile Number ($M = 2.10$, $SD = 2.33$); Duchenne Number ($M = 0.40$, $SD = 0.70$); and Non-Duchenne Number ($M = 1.70$, $SD = 1.89$). For coder D.M., means and standard deviations were: Total Smile Number ($M = 2.90$, $SD = 1.86$); Duchenne Number ($M = 0.80$, $SD = 1.14$); and Non-Duchenne Number ($M = 2.10$, $SD = 1.20$). Similar to the Aim 1 data, D.M.'s coding generated higher scores, as compared to coder A.D.

Reliability of smile Intensity data. As D.M. was the primary coder for smile intensity data, smile intensity reliability was calculated across Times 1, 2, and 3. Intensity reliability was calculated separately for AU6 and AU12. Agreement was defined as both coders having coded the same intensity level, plus or minus 1 intensity level (i.e., if one coder coded level B and the other coder coded level C, this would be categorized as agreement; if, however, one coder coded B and the other coder coded D, this would not count as agreement). A reliability index was then calculated, equaling the number of agreements divided by the sum of agreements and disagreements. For AU6, Intensity agreement was .68. For AU12, Intensity agreement was .95.

Reliability of Onset Duration data. Onset Duration was defined as the time elapsed from the moment of AU12 appearance to the first intensity peak of AU12 (see below for rationale for selecting only AU12 First Peak data for analysis). Reliability of AU12 First Peak coding was defined in terms of the number of agreement divided by the sum of agreements and disagreements. Agreement was defined as both coders identifying the First peak within a specified tolerance window. Agreement was .78 with a ±0.50 second tolerance window and .58 with a ±0.25 second tolerance window. D.M. was coded the First Peak data to be entered into analyses; M.S. coded a subset for the purposes of establishing inter-coder reliability.

Due to the potential impact of inter-coder differences between onset coding Time and 1 and 2 (FACS coded by A.D.) as compared to Time 3 (FACS coded by D.M.), onset data were extracted only from Times 1 and 2.

Data Preparation and Evaluation of Assumptions of Parametric Tests

Missing Data. All subjects were available for baseline and post-treatment assessments but several subjects (n = 3) were not available at six-month follow-up (see Table 1).

Within a number of monologues, temporal data (First Peak and/or Highest Peak) could not be extracted for one of two reasons: 1.) no smiles were produced within the monologue, thus peak intensity data were not available for Duchenne and non-Duchenne smiles, or 2.) a smile was present within a monologue but the onset and/or apex of the smile occurred outside of the one-minute time-frame examined in this study. Due to the limited availability of Duchenne smile peak data, temporal data was averaged across Duchenne and Non-Duchenne smiles; due to the pattern of Highest Peak distribution (see Outliers section, below), temporal analysis was restricted to time-to-apex generated from First Peak measurements. Thus, temporal data were restricted to time-to-apex of AU12s (i.e., First Peak of AU12 minus time of onset of AU12),

averaged across Duchenne and non-Duchenne smiles. For both smile frequency and smile intensity data, if no smiles were present within a monologue, a score of zero was coded.

Outliers in Untransformed Dependent Variables. The presence of outliers, along with the results of normality assessment,arereported in Table 6 andTable 8.A description of the outlier detection process and a summary and discussion of outliers are provided below.

Prior to statistical analysis, data were examined for outliers. Visual inspection of Aim 1 raw numeric data (Lability, Variability, and Complexity scores) did not reveal any evidence of outliers resulting from coder error (e.g., data entry or calculation error). Data were then examined for outliers using Boxplots.[13]

A boxplot of Aim 1 baseline data, grouped by Group (Healthy Control and PD), revealed outliers within Complexity scores among Healthy Controls and within Lability and Variability scores for individuals with PD.

Boxplots of each Aim 1 DV were then examined at each time point (i.e., Baseline [Time 1], Post [Time 2], and 6-Month Follow-Up [Time 3]) grouped by treatment group (Healthy Controls, PD-Untreated, Artic, and LSVT): for Lability, at least one outlier was present in at least one time point in all treatment groups with the exception of LSVT; for Variability, at least one outlier was present in at least one time point for all treatment conditions with the exception of Healthy Controls; for Complexity, at least one outlier was present in at least one time point for all treatment conditions with the exception of PD-Untreated. Because outliers appeared to

[13] Outliers were identified using boxplots in SPSS. The default settings were used, such that a score was identified as an outlier if its value was either 1.) less than or equal to the first quartile minus 1.5 times the interquartile range, or 2.) greater than or equal to the third quartile plus 1.5 times the interquartile range. A description of the SPSS boxplot procedure can be found at http://publib.boulder.ibm.com/infocenter/spssstat/v20r0m0/index.jsp?topic=%2Fcom.ibm.spss.statistics.help%2Fidh_gxss.htm.

represent valid data points (i.e., did appear to result from coding or calculation error), they were not excluded from the data set.

Within the Aim 2 DVs, Time-to-Apex data (i.e., time of First Peak minus time of onset; time of Highest Peak minus time of onset), a number of Highest Peaks were sufficiently temporally distant from smile onset (see, e.g., Ekman & Friesen, 1982) as to suggest that these intensity peaks might be components of subsequent, though not separately coded, smile events. Upon visual inspection of these data, a subset of Highest Peaks appeared to be unrelated to the initial smile onset but, rather, appeared to be related to subsequent smile events (e.g., a gradual increase to a First Peak of B-level intensity might occur shortly after the appearance of a smile, whereas, sometime later, following an intensity plateau and/or decrement, a Highest Peak of C- or D-level intensity might be observed). Because such Highest Peak time points, when compared to the onset of the initial event, could not provide meaningful information about the speed of smile execution—the question of interest within this investigation—the Highest Peak metric was excluded from analyses. All data were judged to be valid within the frequency and intensity DVs.

Each Aim 2 DV was examined for outliers using boxplots. For smile frequency data (Duchenne Number, Non-Duchenne Number, Total Smile Number, and Duchenne Index), examination of a boxplot revealed, at baseline, no outliers within Healthy Controls but multiple outliers for Duchenne Number and one outlier for Total Smile Number. Boxplots were then examined by Treatment Condition at each time point. For Duchenne Number, multiple treatment conditions (PD-Untreated, Artic, and LSVT) had at least one outlier at each time point; for Non-Duchenne Number, outliers were present at multiple time points within the PD-Untreated and LSVT groups; for Total Smile Number, outliers were present at at least one time point within

PD-Untreated, Artic, and LSVT groups; for Duchenne Index, outliers were identified within the Healthy Controls and LSVT groups, at at least one time point.

For intensity data (Intensity of AU6 [within or outside context of a smile], Intensity of Duchenne AU12s, Intensity of Non-Duchenne AU12s, and Overall Smile Intensity), examination of a boxplot revealed outliers, at baseline, among both Healthy Controls (Intensity of Non-Duchenne AU12s and Overall Smile Intensity) and individuals with PD (Intensity of Duchenne AU12s). Boxplots were then examined by Treatment Condition at each time point. For Intensity of AU6, outliers were present at at least one time point for PD-Untreated, Artic, and LSVT groups. For Intensity of Duchenne AU12s, outliers were present at at least one time point for PD-Untreated, Artic, and LSVT. For Intensity of Non-Duchenne AU12s, there was only a single outlier (Healthy Control at baseline). For Overall Smile Intensity, outliers were present at at least one time point within Healthy Controls and LSVT.

For temporal data (First Peak minus onset of AU12), a boxplot of baseline data, grouped by Group, revealed one outlier in both Healthy Controls and in individuals with PD. When examined at both baseline and post-treatment (follow-up data were not collected for this variable – see Methods), three outliers were identified (one within Healthy Controls at baseline, one in ARTIC at baseline, and one in LSVT at post-treatment).

If a dependent variable significantly deviated from normality, data were transformed and re-examined for outliers (see below).Outliers, as defined above, were judged to represent valid data points, did not arise from a single or small number of subjects, and represented a substantial portion of the current data set. Therefore, they were not excluded from analysis.

Distribution of Dependent Variables

All dependent variables were assessed for normality using the Shapiro-Wilk test (Shapiro & Wilk, 1965). If a dependent variable significantly deviated from normality, data were transformed (see below for details) and normality was re-assessed. For outcome data, raw or transformed, demonstrating normality (i.e., not significantly deviating from normality), additional parametric assumptions were evaluated. Results are described below, organized by research aims. For both Aim 1 and Aim 2, DVs were examined by Group (Healthy Control and PD participants) at baseline, and, separately, by treatment group (Healthy Controls, PD-Untreated, PD-Artic, and PD-LSVT) at each time point (Baseline, Post-Treatment, and Follow-Up). Normality assessment findings are reported in Table 6, Table 7, Table 8, and Table 9.

Aim 1a—global measures of facial expressivity among healthy controls andindividuals with PD at baseline. To determine if parametric statistics could be used to evaluate baseline differences between Healthy Controls and individuals with PD, normality was first assessed at Time 1. At Time 1, for each of the Aim 1 DVs—Lability, Complexity, and Variability—the data were not normally distributed within at least one of the comparison groups (i.e., Healthy Controls and/or individuals with PD). Due to significant deviations from normality among raw scores of Healthy Controls and PD participants at baseline, attempts were made to normalize the data via square root, logarithmic (log10), and inverse transformations. Lability, Variability, and Complexity data exhibited positive skew within Healthy Controls and individuals with PD.

Following square root transformation, normal distributions were observed for both Lability and Variability (though, for PD participants, this was marginally the case [$p = .051$]). Following logarithmic transformation, Lability maintained normality; Variability was non-normal within the PD group. For baseline Complexity scores, however, neither square root,

logarithmic, nor inverse transformation yielded normal distributions for both Healthy Controls and PD data.

For square root transformed Lability and Variability scores, the Levene test indicated homogeneity of variance between Healthy Controls and individuals with PD. Boxplot examination of square root transformed baseline Lability data revealed outliers in both Healthy Controls individuals with PD; no outliers were present for baseline Variability distributions.

For log10-transformed Lability scores, homogeneity of variance was demonstrated by the Levene Test, and box-plot inspection revealed one outlier within the Healthy Control group.

Aim 1b—impact of treatment upon global measures of facial expressivity. To determine if parametric procedures could be used to evaluate treatment effects, normality was examined by Treatment Group (Healthy Controls, PD-No Treatment, Artic, and LSVT)) and Time (Time 1 [Baseline], Time 2 [Post-Treatment], and Time 3 [6-Month Follow-Up]. For Lability, Variability, and Complexity, raw score distributions significantly deviated from normality within a number of Treatment Condition × Time cells.

Following square root transformation: Variability scores deviated from normality within the LSVT group at Time 3; Lability scores deviated from normality within the Untreated PD group at both Time 1 and Time 2; and Complexity scores deviated from normality across a number of Treatment Group × Time combinations. Because the square root transformed Variability scores deviated from normality in only one cell, the data were assessed for homogeneity of variance and outliers. The square-root-transformed Variability data exhibited homogeneity of variance as assessed by Levene's Test for Equality of Variances. Inspection of box-plots revealed outliers within the PD-Untreated and LSVT groups at Time 3. For each Aim 1 DV, additional transformations were examined for normality.

Following logarithmic transformation (log10), Lability scores significantly deviated from normality only within the Healthy Controls at Time 2; Variability exhibited normality across all Treatment Group × Time combinations, though this was marginally the case for the LSVT group at Time 1; Complexity scores significantly deviated from normality within multiple cells. Log10-transformed Lability scores exhibited homogeneity of variance as assessed by Levene's Test for Equality of Variances; inspection of box-plots revealed multiple outliers within the Healthy Control group and a single outlier in the PD-Untreated group. Log10-transformed Variability scores also exhibited homogeneity of variance as assessed by Levene's Test for Equality of Variances; inspection of box-plots revealed a single outlier (LSVT group at Time 2).

Following inverse transformation, Lability scores significantly deviated from normality across all Treatment Group × Time combinations; Variability scores significantly deviated from normality across most Treatment Group × Time combinations; and Complexity scores significantly deviated from normality across a number of Treatment Group × Time combinations.

Distribution of Aim 1 change-scores. Distributions of Time 3 minus Time 1 change-scores were examined. Although some variables demonstrated normality and homogeneity of variance, each transformed distribution continued to exhibit outliers.

For Lability, change-scores lacked normality and most treatment groups exhibited positive skew. Square root, logarithmic, and inverse transformations failed to normalize the data.

For Variability, change-scores lacked normality and exhibited positive skew across all treatment groups. Square root transformation normalized the Variability change-score distribution, and demonstrated homogeneity of variance across treatment groups, but

examination of box-plots revealed multiple outliers in multiple treatment groups. Logarithmic transformation also normalized the data, exhibited homogeneity of variance, but again, examination of box-plots revealed multiple outliers in multiple treatment groups. Inverse transformation failed to normalize Variability change-score data.

For Complexity, raw change-scores exhibited negative skew in most treatment groups and significantly deviated from normality only within the Healthy Control group; outliers were present only within the Healthy Control group. Transformations were modified to account for negative skew. Square root transformation failed to normalize the data. Logarithmic transformation normalized the data and exhibited homogeneity of variance, but examination of box-plots revealed two outliers within the Healthy Control groups. Inverse transformation yielded normalized data and homogeneity of variance, but box-plot examination revealed multiple outliers (one in ARTIC and one in LSVT).

Distributions of Aim 1 Time 3 minus Time 2 changes-scores were examined. For Lability, change-scores lacked normality within two of the four treatment groups. Three of four treatment groups exhibited positive skew. For Variability, Time 3 minus Time 2 change-scores lacked normality within two of the four treatment groups; all groups exhibited positive skew. For Complexity, Time 3 minus Time 2 change-scores lacked normality within one of the four treatment groups; each group exhibited negative skew. As the above T3-T1 change-score transformations failed to normalize the data and/or adequately manage homogeneity of variance and/or outliers, nonparametric statistics were selected for analyses; accordingly, transformations were not conducted on the T2-T1 change-scores.

Aim 2a—smiling behavior inhealthy controls and individuals with PD at baseline.

The dependent variables used to investigate the primary and exploratory hypotheses concerning smiling behavior concern three aspects of smile behavior: frequency (Number of Duchenne Smiles, Number of Non-Duchenne Smiles, Total Number of Smiles [Duchenne plus non-Duchenne], and Duchenne Index (DI; DI = (D-ND)/(D+ND)), intensity (AU6s [alone or in combination with AU12], AU12s in Duchenne smiles, AU12s non-Duchenne smiles, and overall intensity of AU12s), and temporal dynamics (AU12 First Peak minus Onset).

To determine if parametric statistics could be used to evaluate baseline differences between Healthy Controls and individuals with PD, DV normality was first assessed at Time 1; to determine if parametric statistics could be used to evaluate treatment effects, normality was assessed for each DV, in each treatment group, at each time point. Normality data are organized by type of DV (i.e., frequency, intensity, or temporal).

Frequency. At baseline, each frequency DV exhibited statistically significant departure from normality within the Healthy Control and/or PD groups. In light of the statistically significant departures from normality, attempts were made to normalize the data via transformation of raw scores. To determine which transformations might be appropriate, raw-score distributions were examined. Deviations from normality appeared to result, at least in part, from a large number of subjects who did not produce smiles, thus resulting in a large number of zero scores at one end of the score distribution. Almost all frequency DVs by Group (HC or PD) cells exhibited positive skew. Square root, logarithmic, and inverse transformations were conducted. Following square root[14], $\log_{10}$[15], and inverse[16] transformation, each variable

[14] Because the square root operation requires positive scores, the absolute value of the largest negative value in the untransformed Duchenne Index data, -2, was added to the raw scores. For negatively skewed DVs, prior to applying the square root function, raw scores were reflected (i.e., resulting in the formula: $\sqrt{}$ ((largest score in the distribution + 1)-score)).

continued to exhibit a significant deviation from normality in either one or both Groups (HC or PD).

Intensity. At baseline, each intensity DV exhibited statistically significant departure from normality within one or both Groups (HC or PD; see Table 7). In light of the statistically significant departures from normality, attempts were made to normalize the data via transformation of raw scores. To determine which transformations might be appropriate, raw-score distributions were examined. At baseline, AU6 intensity and AU12-Duchenne intensity exhibited positive skew within both Healthy Control and PD groups; AU12-Non-Duchenne and AU12-Overall intensity exhibited both positive and negative skew. In light of these distribution characteristics, square root, logarithmic, and inverse transformations were performed on all intensity variables; additionally, for AU12-Non-Duchenne and AU12-Overall, each transformation was again performed, with the transformation formula modified to correct for negative skew[17]. For each smile intensity DV, square root, $\log_{10}$, and inverse transformations to

---

[15] Logarithmic transformations can only be applied to positive, non-zero, data. For positively skewed DVs, prior to log transformation, a constant of +1 was added to Duchenne, Non-Duchenne, and Total Smile scores; a constant of +3 was added to Duchenne Index scores; a constant of +1 was added to all smile intensity DVs. For negatively skewed DVs, prior to logarithmic transformation, raw scores were reflected (i.e., resulting in the formula $\log_{10}$((largest score in the distribution + 1)-score)).

[16] Inverse transformations can only be applied to non-zero data. To remove zeros, a constant of +1 was added to Duchenne, Non-Duchenne, and Total Smile scores; a constant of +3 was added to Duchenne Index scores. To maintain the relative positions of individual scores, the inverse transformation took the form of -1/(score + constant). For positively skewed intensity DVs, a constant of +1 was added; for negatively skewed intensity DVs, raw scores were reflected (i.e., resulting in the formula 1/((largest score in the distribution + 1)-score)).

[17] For variables exhibiting negative skew, the following formulas were employed. . For negatively skewed DVs, prior to applying the square root function, raw scores were reflected (i.e., resulting in the formula sqrt((largest score in the distribution + 1)-score). For negatively skewed DVs, prior to logarithmic transformation, raw scores were reflected (i.e., resulting in the formula $\log_{10}$((largest score in the distribution + 1)-score)). For negatively skewed intensity DVs, prior

correct for positive skew failed to yield normality within the Healthy Control and/or PD groups. For both non-Duchenne AU12 intensity and overall AU12 intensity at baseline, negative skew-corrected transformations (square root and logarithmic) also failed to yield normality in either one or both of the Healthy Control and PD groups.

Temporal. Raw baseline temporal data (AU12: First Peak minus Onset) exhibited normality within the Healthy Control group but significantly deviated from normality in individuals with PD ($p = .03$) both groups were positively skewed. Square root, logarithmic (log10), and inverse transformations were performed in an effort to normalize the data.

Following square root transformation, both Healthy Controls and the PD group exhibited normality. Examination of box-plots revealed no outliers in either group; data exhibited homogeneity of variance as assessed by Levene's Test for Equality of Variance.

Following logarithmic transformationboth Healthy Controls and the PD group exhibited normality. Examination of box-plots revealed no outliers in either group, at baseline; data exhibited homogeneity of variance as assessed by Levene's Test for Equality of Variance.

Following inverse transformation, the PD group lacked normality and inspection of box-plots revealed outliers.

Aim 2b—smiling behavior after treatment.

To determine if parametric procedures could be used to evaluate treatment effects upon smiling behavior, normality of each DV was examined within each Treatment Group (Healthy Controls, PD-Untreated, Artic, and LSVT)) at each Time (Time 1 [Baseline], Time 2 [Post-Treatment], and Time 3 [6-Month Follow-Up]) cell. Change-scores were also examined. As Time 2 minus Time 1 change-scores lacked normality and/or homogeneity of variance, and/or

---

to calculating an inverse, raw scores were reflected (i.e., resulting in the formula 1/((largest score in the distribution + 1)-score)).

exhibited outliers, even after transformations, nonparametric statistics were selected for analyses; accordingly, transformations were not conducted on Time 3 minus Time 1 change-scores or Time 3 minus Time 2 change-scores. Normality testing and the results of transformation are summarized below.

Frequency. Duchenne Number deviated from normality across all treatment groups and times; all Treatment Group x Time cells but one exhibited positive skew. Square root, logarithmic, and inverse transformations failed to normalize the data. Non-Duchenne Number data deviated from normality within multiple treatment groups at each time point and exhibited positive skew across Treatment Group × Time cells. Square root logarithmic, and inverse transformations failed to normalize the data. Total Smile Number data deviated from normality within multiple treatment groups at each time point; data exhibited positive skew across all but one treatment group x time cell. Square root, logarithmic, and inverse transformations failed to normalize the data. Duchenne Index data deviated from normality across most treatment group x time cells and exhibited positive skew across most Treatment Group × Time cells.Square root, logarithmic, and inverse transformations failed to normalize the data.

Intensity. AU6 intensity data deviated from normality across most Treatment Group × Time cells; data exhibited positive skew within most treatment group x time cells. Square root, logarithmic, and inverse transformations failed to normalize the data. AU12-Duchenne intensity data deviated from normality within all but one treatment group x time cell and exhibited positive skew in all but one Treatment Group × Time cells. Square root, logarithmic, and inverse transformations failed to normalize the data. AU12-Non-Duchenne intensity deviated from normality within multiple treatment groups at each time point; most treatment group x time cells exhibited negative skew. Negative skew-adjusted square root, logarithmic, and inverse

transformations failed to normalize the data. AU12-Overall intensity deviated from normality with multiple treatment groups at each time point (see Table 9); most Treatment Group × Time cells exhibited negative skew. Negative skew-adjusted square root, logarithmic, and inverse transformations failed to normalize the data.

Temporal. When examined by treatment group at each time point (only baseline and post-treatment for temporal data), Temporal data (AU12 first peak minus onset) deviated from normality within ARTIC at baseline and within the LSVT group at post-treatment; most treatment group × time cells exhibited positive skew. Following square root transformation, the LSVT group, at post-treatment, still lacked normality; groups exhibited homogeneity of variance and box-plot inspection revealed no outliers. Following logarithmic transformation, all cells demonstrated normality, exhibited homogeneity of variance, and inspection of a box-plot revealed one outlier within the ARTIC group post-treatment. Following inverse transformation, multiple cells lacked normality.

Distribution of Frequency change-scores.Duchenne Number change-scores (Time 2 minus Time 1) lacked normality and exhibited both positive and negative skew across different treatment groups (highest skew values were positive). Square root, logarithmic, and inverse, transformation failed to normalize the data.

Non-Duchenne Number change-scores exhibited normality and homogeneity of variance, but inspection of box-plots revealed multiple outliers within the PD-Untreated group. Non-Duchenne change-scores exhibited both positive and negative skew; highest skew values were negative. Square root, logarithmic, and inverse transformations failed to normalize the Non-Duchenne change-score data.

Total Smile Number change-scores exhibited normality and homogeneity of variance, but box-plot inspection revealed multiple outliers across multiple treatment groups. Skew was negative in most treatment groups. Square root transformed scores exhibited normality and homogeneity of variance, but box-plot examination revealed multiple outliers in multiple treatment groups. Logarithmic transformation resulted in normal distribution and homogeneity of variance, but box-plot examination revealed multiple outliers within multiple treatment groups. Inverse transformation did not result in normalization.

Duchenne Index change-scores lacked normality; skew was positive in most treatment groups. Square root, logarithmic, and inverse transformation failed to normalize the Duchenne Index change-score data.

Distribution of Intensity change-scores. AU6 Intensity change-scores significantly deviated from normality within multiple treatment groups; change-scores were negatively skewed. Square root, logarithmic, and inverse transformations failed to normalize AU6 Intensity change-scores. AU12-Duchenne Intensity change-scores lacked normality; skew was negative within most treatment groups. Square root, logarithmic, and inverse transformations failed to normalize AU12-Duchenne Intensity change-score data. AU12-Non-Duchenne Intensity change-scores exhibited normality and homogeneity of variance, but box-plot inspection revealed outliers in multiple treatment groups; skew was negative within most treatment groups. Square root transformation yielded normality but lacked homogeneity of variance, and outliers remained in multiple treatment groups. Logarithmic transformation yielded normality but lacked homogeneity of variance, and multiple outliers remained. Inverse transformed AU12-Non-Duchenne Intensity data lacked normality in multiple treatment groups. AU12-All Intensity change-scores exhibited normality and homogeneity of variance, but two outliers were present in

the ARTIC group. Treatment groups exhibited positive and negative skew. Two sets of transformations were conducted: one for negative skew and one for positive skew. With negative skew-adjusted transformations, square root, logarithmic, and inverse transformations failed to normalize the AU12-All Intensity change-score data. With positive skew-adjustment, square root, logarithmic, and inverse transformations also failed to yield normality.

Distribution of Temporal change-scores. For temporal data (AU12 First Peak minus onset), change-scores significantly deviated from normality within the Healthy Control group and box-plot examination revealed multiple outliers. As positive and negative skew was evident in different treatment groups, two sets of transformations were conducted. For positive skew-correcting transformations: square root transformation lacked normality in the Healthy Control group, exhibited homogeneity of variance, and box-plot inspection revealed multiple outliers; logarithmic transformation also lacked normality in the Healthy Control group, exhibited homogeneity of variance, and box-plot inspection revealed multiple outliers; inverse transformation lacked normality in multiple treatment groups. For negative skew-correcting transformations: square root transformation lacked normality in Healthy Controls, exhibited homogeneity of variance across treatment groups, and inspection of box-plots revealed multiple outliers; logarithmic and inverse transformation was non-normal in multiple groups.

Distribution of Demographic Variables

As some analyses incorporated the Beck Depression Inventory (BDI), BDIdistributions were examined. When examined by Group, BDI scores at Time 1 (baseline) deviated significantly from the normal distribution for both Healthy Controls ($p = .03$) and individuals with PD ($p = .02$). As a natural log transformation was shown to normalize BDI data within the same data set (Alterescu, 2012), a natural log transformation was performed. When re-examined

by Group, neither group significantly deviated from the normal distribution (HC: $p = .12$. PD: $p = .05$) but this was marginally the case for the PD group. When examined by treatment group, no group significantly deviated from the normal distribution.

Other demographic variables were also examined. When examined by Group (HC or PD), Age was normally distributed across both groups, but Education and MMSE were not. When examined by treatment group, Age was normally distributed across groups, as was Years of Education, but MMSE score was not.

With regard to PD-specific variables, neither Time Since Diagnosis nor the Hoehn and Yahr stage (Hoehn &Yahr, 1997) were normality distributed across the three PD groups; for both variables, two of three PD treatment groups significantly deviated from normality.

Findings

Organization of findings. Statistical analyses examining baseline differences between Healthy Controls (HCs) and individuals with Parkinson's disease (PDs), as well as treatment effects (i.e., as a function of the four Treatment groups: HC, PD-Untreated, ARTIC, and LSVT), are reported below. First, findings related to global measures of facial expressivity (Lability, Variability, and Complexity) are reported. Following this, smile behavior findings are reported; smile behavior variables are discussed under the subheadings of Frequency, Intensity, and Onset Duration. In both sections (global measures and smile behavior), baseline findings are discussed first and are followed by an examination of treatment effects. For most analyses, nonparametric methods were selected in light of data not meeting assumptions of parametric methods. Specific analyses are described in greater detail within in each subsection.

Gender. It is wellestablished that women tend to be more emotionally expressive than

men (e.g., Borod & Madigan, 2000; Grunwald et al., 1999; LaFrance et al., 2003). Within the current data set, there is an imbalance of men and women in the Healthy Control group (63.6% women) relative to the PD group (26.7% women).Based on this information, we considered the possibility that anyoverall group (HC versus PD) differences could result entirely or in part from the predominance of women in the Healthy Control group and the predominance of men in the PD group.

To better understand the impact of gender within the current data set, each baseline variable was examined for gender differences (see Table 12) in both the HC and PD groups.For each variable with the exception of Onset Duration, gender effects were assessed with a Mann-Whitney test (Onset Duration gender effects were assessed with a*t*-test; all variables with the exception of Onset Duration were assessed with a one-tailed tests).In light of the relatively small sample sizes, results with *p*-values of .20 or smaller are described below. Within the PD group, of the 12 variables tested for gender differences, statistically significant ($p \leq .05$) gender differences were obtained for 4 variables (Lability, Variability, Complexity, and Duchenne Number), 2 variables had *p*-values between .05 and .10 (Duchenne AU12 Intensity and Overall AU12 Intensity), and 2 variables had *p*-values between .10 and .20 (AU6 Intensity and Onset Duration). For each of these potential gender differences, with the exception of Onset Duration, women produced higher scores than men. Within the Healthy Control group, of the 12 variables tested for gender differences, 2 variables demonstrated gender differences at the $p \leq .05$ level (Non-Duchenne AU12 Intensity and Overall AU12 Intensity), 3 variables had *p*-values between .05 and .10 (Duchenne Number, AU6 Intensity, and Duchenne AU12 Intensity), and no variable had*p*-value between .10 and .20; for each of these potential gender differences, women exhibited higher scores relative to men. In summary, gender differences were observed for some variables,

especially within the PD group, and when potential gender differences were found, women typically exhibited higher levels of expressivity as compared to men.

To address the potential gender confound in comparing baseline HC and PD groups, overall group findings are reported first, and are followed by separate group analyses for men and women.Gender-specific results, as compared to overall results,are necessarily calculatedwith smaller samples. Thus, they may suffer from limited statistical power. Nevertheless, examination of gender-specific findings canhelp determine if the overall group findings are attributable to group characteristics per se (i.e., expressivity differences between healthy controls and individuals with PD) or, whether the observed group differences might be an artifact of the larger number of women in the Healthy Control group. If the results are entirely attributable to greater expressivity among healthy individuals, as compared to people with PD, the same pattern of results should hold when examined separately by gender. To the extent that the group results were obtained due to the high proportion of women (who, as a group, tend to be more emotionally expressive than men) in the Healthy Control group), HC/PD differences will not exist when examined separately by gender. Due to limitations of sample size, gender-specific analyses are restricted to baseline analyses (which compare two groups) and are not applied to treatment analyses (which compare four) groups.

Global measures of facial expressivity.

Baseline global expressivity differences between Healthy Controls and individuals with PD. To provide context for the current follow-up (Time 3) findings, Dumer's (2011) baseline (Time 1) and post-treatment (Time 2) results are briefly summarized below. Dumer (2011) demonstrated, via a Welch-James procedure, that baseline Lability in Healthy Controls (M = 15.27, SD = 10.98) was significantly greater than in individuals with PD (M = 7.27, SD = 7.35),

$F(1, 12.34) = 5.36$, $p = .04$, $ES = 1.15$. Additionally, the mean Variability of Healthy Controls ($M = 7.00$, $SD = 4.77$) was marginally significantly greater than the Variability of individuals with PD ($M = 3.80$, $SD = 2.79$), $F(1, 12.09) = 4.50$, $p = .055$, $ES = 0.85$. Complexity however did not significantly differ, $F(1, 13.13) = 1.22$, $p = .29$, between controls ($M = 1.83$, $SD = 0.97$) and individuals with PD ($M = 1.51$, $SD = 0.79$). As an exploratory measure, to ensure comparability with the nonparametric tests employed in the current study, these data were reexamined with Mann-Whitney U Tests (see Table 15): again, there was evidence that Healthy Controls exhibited higher Lability, $U = 122.50$, $z = -2.59$, $p = .01$, and Variability scores, $U = 145.50$, $z = -2.12$, $p = .03$, but not Complexity scores, $U = 215.50$, $z = -0.66$, $p = .51$. These analyses demonstrate pre-treatment differences between PD and non-PD participants for Lability and Variability.Notably, examined separately by Gender, there were no statistically significant differences between HC ($n = 7$) and PD ($n = 12$) groups among women, though for each of the three variables, the group mean was higher for Healthy Controls as compared to the PD group. Among men, HCs ($n = 4$) exhibited significantly more Lability and Variability, compared to PDs ($n = 33$). There was no statistically significant difference between groups for Complexity. See Table 15.

Global measures of facial expressivity: Baseline to post-treatment (Time 2). In the Emotion Lab's analysis of post-treatment effects (Dumer, 2011) employed separate Welch-James procedures for each of the three measures of facial expressivity (Lability, Variability, and Complexity) to examine differences in change-scores (Time 2 minus Time 1) between the LSVT treatment group and the combined non-LSVT groups (i.e., comprised of Healthy Controls, PD-No Treatment, and PD-Artic). The LSVT group, compared to non-LSVT groups, was found to have significantly greater increases in Variability and Lability, but not Complexity. Dumer

(2011) also directly compared the change-scores between the LSVT and ARTIC groups: the degree of Variability change was marginally significant whereas LSVT and ARTIC did not significantly differ in the degree of change of Lability and Complexity change-scores.

The Time 1 to Time 2 change-score findings assess treatment effects through detection of between-group differences in degree of treatment-related change. To complement the change-score findings—and provide a direct measure of within-group change—Time 1 Lability, Variability, and Complexity scores, were directly compared against Time 2 scores using separate Wilcoxon signed-rank tests for each of the four Treatment Groups.

For both Lability and Variability, the LSVT group exhibited higher scores at Time2, as compared to Time 1 ($z = 2.24$, $p = .03$; $z = 2.11$, $p = .04$); no other treatment group exhibited statistically significant changes in Lability or Variability when comparing baseline to post-treatment. For Complexity, no treatment groups exhibited statistically significant changes from baseline to post-treatment, but, at the trend-level, Healthy Controls did exhibit higher scores at Time 2, $z = 1.89$, $p = .06$.

Due to violations of parametric assumptions of overall facial expressivity data (Lability, Variability, and Complexity)—for each individual Time (1, 2, and 3) and for the three change-scores, the facial expression data were analyzed via change-scores, using the nonparametric one-way Kruskal-Wallis analysis of variance by ranks test. Accordingly, the pre-/post-treatment findings analyzed by Dumer (2011) were re-examined using the one-way Kruskal-Wallis procedure to facilitate interpretation of the follow-up data.

For each Kruskal-Wallis test conducted, the dependent variable was the relevant Aim 1 change-score (i.e., Lability, Variability, or Complexity) and Treatment Group (Healthy Control, PD-Untreated, ARTIC, and LSVT) was the between-subjects factor. Analysis of T2-T1 change-

scores was conducted separately for Lability, Variability, and Complexity, revealing significant treatment group differences for Variability, $\chi^2(3) = 8.01$, $p = .05$, a trend for Lability, $\chi^2(3) = 7.28$, $p = .06$, and no statistically significant treatment group differences for Complexity $\chi^2(3) = 4.21$, $p = .24$. Pairwise comparisons are described below. Means, standard deviations, and medians of T2-T1 change-scores are reported in Table 13.

For Variability, planned pairwise comparisons, with no alpha adjustment, revealed a significant difference in change-scores between the LSVT group and the PD-Untreated group (p=.01), such that from baseline to post-treatment, LSVT (M = 1.81; Mdn = 1.50) increased in Variability, relative to the PD-Untreated group (M = -0.94; Mdn = -1.00). Although a planned pairwise comparison of Variability T2-T1 change-scores in the LSVT versus ARTIC groups did not significantly differ with alpha set at .05, a trend-level difference was observed (p = .08) such that, from baseline to post-treatment, the LSVT group, relative to Artic, had a greater increase in Variability. Upon examination of the remaining pairwise comparisons, LSVT change-scores were greater than Healthy Control change-scores (p = .04); no other treatment group differences in T2-T1 Variability change-scores were observed. Means, standard deviations, and medians of T2-T1 change-scores are reported in Table 13.

Although treatment group differences between Lability T2-T1 change-scores were not significant at the alpha level of .05, planned pairwise comparisons, without alpha adjustment, were conducted to explore the observed trend(p = .06). The LSVT group significantly differed from the PD-Untreated group (p = .01) such that LSVT T2-T1 change-scores (M = 5.00; Mdn = 2.50) exceeded those of PD-Untreated (M = -1.71; Mdn= -1.00). A trend-level difference was observed (p = .06) such that, from baseline to post-treatment, LSVT Lability change-scores, relative to ARTIC (M = -1.25 ,Mdn = .50), increased more in Lability. On an exploratory basis,

the remaining pairwise comparisons were examined: LSVT change-scores were greater than Healthy Control change-scores (p = .09); no other treatment group differences were observed.

Global measures of facial expressivity at follow-up (Time 3).

To determine if treatment groups differed in their effect at Time 3, separate Kruskal-Wallis (K-W) procedures[18] were performed on each global expressivity measure (Lability, Variability, and Complexity). See Table 16. For each K-W procedure, the dependent variable was a change-score; K-W procedures were conducted separately for Time 3 minus Time 2 change-scores and for Time 3 minus Time 1 change-scores. The independent factor was Treatment Group (Healthy Controls, PD-No Treatment, PD-Artic, and PD-LSVT). For each K-W procedure, if the K-W omnibus statistic was significant, post-hoc Mann-Whitney tests were used to examine group contrasts.

For each of the three overall facial expressivity T3-T1 change-scores, there were no statistically significant differences among the four treatment groups (Lability change-scores (T3-T1):$\chi^2(3) = 3.39$, p =.34. Variability change-scores (T3 minus T1):$\chi^2(3) = 2.12$, p = .55. Complexity change-scores (T3 minus T1): $\chi^2(3) = 2.98$, p = .39).

For Time 3 minus Time 2 change-scores, again, there were not statistically significant differences among treatment groups for any of the global facial expressivity variables: Lability, $\chi^2(3) = 0.53$, p = .91; Variability, $\chi^2(3) = 1.07$, p = .78; Complexity, $\chi^2(3) = 3.89$, p = .27.

Means, standard deviations, and medians, for T2-T1, T3-T2, and T3-T1 change-scores, are reported in Table 13.

Smilebehavior: baseline characteristics of Healthy Control and PD groups

Baseline smile behavior data, comparing Healthy Controls and individuals with PD, are

[18] As the Aim 1 data did not conform to parametric assumptions, non-parametric procedures are employed.

reported below, grouped by variable type (Frequency, Intensity, and Temporal). Each measure of baseline smile behavior was also analyzed separately by Gender. Gender-specific findings are reported following the overall findings. Overall and gender-specific data are summarized, respectively, inTable 17,andTable 18.

Baseline differences in smile frequency between Healthy Controls and individuals with PD. To determine if Healthy Controls and individuals with PD differed in smile frequency at baseline, Mann-Whitney U tests were performed on each of the (untransformed) smile frequency dependent variables (Duchenne Number, Non-Duchenne Number, Total Smile Number, and Duchenne Index). In light of prior research, a one-tailed test was used to evaluate baseline differences in Duchenne Number. All other baseline frequency variables were assessed using two-tailed tests. For Duchenne Number, Healthy Controls and individuals with PD did not significantly differ at the alpha level of .05, but did exhibit a trend-level difference (one-tailed[19]) such that Duchenne Number mean rank was greater in Healthy Controls ($M = 0.73$, $Mdn = 0.00$) as compared to individuals with PD ($M = 0.24$, $Mdn = 0.00$), $U = 194.00$, $z = -1.54$, $p = .06$). For Non-Duchenne Number, Healthy Controls and individuals with PD did not differ at the alpha level of .05 but exhibited a trend-level difference such that the mean rank ofHealthy Controls ($M = 2.36$, $Mdn = 2.00$) was greater than the mean rank of individuals with PD ($M = 1.13$, $Mdn = 1.00$), $U = 158.50$, $z = -1.91$, $p = .06$. The Duchenne Index, equal to the number of Duchenne Smiles minus the number of Non-Duchenne smiles, divided by the sum of Duchenne and Non-Duchenne smiles, did not significantly differ between Healthy Controls and individuals with PD, $U = 244$, $z = -0.08$, $p = .47$). On an exploratory basis, Total Smile Number was also examined;

[19] For Duchenne Number, the one-tailed *p*-value of .06 was calculated by dividing the two-tailed Mann Whitney *p*-value of .124 by 2. All other baseline frequency measures were assessed at the two-tailed level, unless otherwise indicated.

the mean rank of Healthy Controls (M = 3.09, Mdn = 3.00) was significantly greater than the mean rank of individuals with PD (M = 1.38, Mdn = 1.00), U = 133.00, z = -2.44, p = .02.

When smile frequency measures were analyzed separately by Gender, there were no significant differences, or trend-level differences, between HCs (Women: n = 7. Men: n = 4) and PDs (Women: n = 12. Men: n = 33). As can be seen in Table 19,among women, although not statistically significant, the mean number of Duchenne, Non-Duchenne, and Total smiles was greater in HCs as compared to PDs. In men, however, the HC group did not produce any Duchenne smiles, thus the mean of the PD group was higher than the HC group. The Duchenne Index was similar between HC and PD groups for the overall data set, as well as when analyzed separately by Gender.

Baseline differences in smile intensity between Healthy Controls and individuals with PD. To determine if Healthy Controls and individuals with PD differed in smile intensity at baseline, one-tailed[20] Mann-Whitney U tests were performed on each of the smile intensity dependent variables (AU6 Intensity, Duchenne AU12 Intensity, and Non-Duchenne AU12 Intensity). There were no significant differences between Healthy Controls and individuals with PD in AU6 Intensity (U = 229.00, z = -0.47, p = .32) or Duchenne AU12 Intensity (U = 205.00, z = -1.19, p = .12). There was however, a statistically significant difference, U = 160.50, z = -1.85, p = .03, between Healthy Controls and individuals with PD with respect to Non-Duchenne AU12 Intensity, such that the mean rank of the Healthy Control group (M = 2.05, Mdn = 2.33) exceeded that of the PD group (M = 1.31, Mdn = 1.33). On an exploratory basis, Overall AU12 Intensity was examined. Although Overall AU12 Intensity did not significantly differ between Healthy Controls and individuals with PD, U = 175.00, z = -1.54, p = .06, a trend was observed

[20] For smile intensity measures, one-tailed p-values were calculated by running two-tailed Mann Whitney tests in SPSS and dividing the obtained p-value by 2.

such that the mean rank of Healthy Controls (M = 2.19, Mdn = 2.33) exceeded that of the PD group (M = 1.47, Mdn = 2.00).

When the Intensity variables were examined separately for men (HC: n = 4. PD: n = 33) and women (HC: n = 7. PD: n = 12), there were no statistically significant or trend-level differences between HC and PD groups. Among women, although HC and PD groups did not significantly differ for any Intensity variable, the mean values for all Intensity variables were greater in HCs as compared to PDs. Among men, this was true for Non-Duchenne AU12 and Overall AU12 intensity but, as there were no instances of AU6 or Duchenne smiles within the HC group, the mean intensity of these measures were greater in the PD group.

*Baseline differences in smile onset duration between Healthy Controls and individuals with PD.*In light of a potential relationshipbetween smile onset duration and smile intensity, ANCOVA was selected to examine baseline differences in smile onset duration (log10-transformed) while controlling for smile intensity.

Prior to conducting the analysis, ANCOVA assumptions were assessed. As noted above, log10-transformed AU12 Onset Duration met parametric assumptions. Although the overall distribution for AU12 Intensity at baseline (and transformations of this measure) was not normally distributed, the subset used within the ANCOVA analysis did exhibit normality, homogeneity of variance, and did not include outliers.[21] Thus, baseline correlations were calculated using Pearson's r. The overall baseline correlation between log10-transformed AU12

[21] Because, by definition, Onset Duration scores can only exist when an AU exists at > 0 intensity, AU12 Intensity scores of zero are not included in the ANCOVA. Therefore, it was necessary to re-assess the distribution of AU12 Intensity, screening out the zero scores. After screening out zero scores, AU 12 Intensity no longer exhibited a statistically significant departure from normality, as demonstrated by the Shapiro-Wilk test (HCs, p = .20; PDs, p = .24). After screening out zero scores, Levene's test suggested homogeneity of variance across Healthy Controls and individuals with PD p = .81).

Onset Duration and AU12 Intensity was statistically significant (Pearson's $r = .40$, $p = 02$). When examined by Group (HC versus PD), the correlation was not statistically significant for Healthy Controls (Pearson's $r = .40$,$p = .28$) but was statistically significant for the PD group (Pearson's $r = .43$, $p = .04$). Although a statistically significant correlation did not hold for both groups, ANCOVA assumptions were further examined. A Mann Whitney test showed that baseline Overall AU12 Intensity did not significantly differ between Healthy Controls and individuals with PD, $U = 175.00$, $z = -1.54$, $p = .13$ (t-test on the onset duration subset was also not significant, $t(34) = 0.76$, $p = .45$). The assumption of homogeneity of regression slopes was also satisfied, as demonstrated by an ANOVA indicating a statistically insignificant interaction between the independent factor and the covariate (i.e., Group [HC and PD] and Overall AU12 Intensity), $F(1, 28), = .03$, $p = .86$)[22]. Standardized residuals of log10-transformed AU12 Onset Duration, for both Groups (HC and PD) and for the overall model, were normally distributed, as demonstrated by the Shapiro-Wilks test ($p > .05$). There was homoscedasticity and homogeneity of variances, as assessed by visual inspection of a scatterplot and Levene's Test of Homogeneity of Variance ($p = .77$), respectively. There was a single outlier in the data, as assessed by cases with standardized residuals greater than ±3 standard deviations. Thus, with the caveats outlined above, the assumptions of the ANCOVA were met to the extent that ANCOVA analysis seemed justified. An ANCOVA was performed both with and without the single outlier described above.

An ANCOVA was performed with log10-transformed AU12 Onset Duration as the dependent variable, AU12 Intensity as the covariate, and Group (Healthy Controls and individuals with PD) as the independent factor. With the single outlier included, after covarying for baseline AU12 Intensity, there was not a statistically significant difference in log10-

[22] Because the covariate was not normally distributed, the use of an ANOVA-based test may have failed to detect an interaction.

transformed AU Onset Duration between the Healthy Controls and individuals with PD$F(1,29) = .76$, $p = .39$, partial $\eta^2 = .03$. With the outlier excluded, after covarying for baseline AU12 Intensity, there was not a statistically significant difference in log10-transformed AU Onset Duration between the Healthy Controls and individuals with PD$F(1,28) = 2.33$, $p = .14$, partial $\eta^2 = .08$. Estimated marginal means for Healthy Controls and PD groups were, respectively, -0.44 and -0.29,[23] indicated that individuals with PD had longer smile onset durations, even after controlling for smile intensity.

On an exploratory basis, a separate ANCOVA was performed, this time using log10-transformed AU12 Intensity scores as a covariate. ANCOVA assumptions were assessed. Again, both the dependent variable, as well as the covariatedistributions,met parametric assumptions.Overall baseline correlation between log10-transformed AU12 Onset Duration and log10-transormed AU12 Intensity are described above (as non-parametric correlations were calculated, these values are unchanged after transformation of the covariate). A Mann-Whitney test demonstrated that baseline Overall AU12 Intensity did not significantly differ between Healthy Controls and individuals with PD. The assumption of homogeneity of regression slopes was also satisfied, as demonstrated by an ANOVA indicating a statistically insignificant interaction between the independent factor and the covariate (i.e., Group [HC and PD] and Overall AU12 Intensity), $F(1,28) = <.01$, $p = .97$). Standardized residuals of log10-transformed AU12 Onset Duration, for both Groups [HC and PD] and for the overall model, were normally distributed, as demonstrated by the Shapiro-Wilks test ($p > .05$). There was homoscedasticity and homogeneity of variances, as assessed by visual inspection of a scatterplot and Levene's Test of Homogeneity of Variance ($p = .75$), respectively. There was a single outlier in the data, as

[23] Values are negative due to log10-transformation of fractions. More positive values indicate longer onset duration.

assessed by cases with standardized residuals greater than ±3 standard deviations. With the caveats outlined above, the assumptions of the ANCOVA were met to the extent that ANCOVA analysis seemed justified. The ANCOVA was run with and without the outlier.

An ANCOVA was performed with log10-transformed AU12 Onset Duration as the dependent variable, log10-transformed AU12 Intensity as the covariate, and Group (HC and PD) as the independent factor. With the single outlier included, after covarying for baseline log10-transformed AU12 Intensity, there was not a statistically significant difference in log10-transformed AU Onset Duration between the Healthy Controls and individuals with PD$F(1,29) = .81$, $p = .38$, partial $\eta^2 = .03$. With the outlier removed and covarying for baseline log10-transformed AU12 Intensity, the data, again, did not exhibit a significant difference in log10-transformed AU Onset Duration between Healthy Controls and individuals with PD $F(1,28) = 2.49$, $p = .13$, partial $\eta^2 = .08$. Estimated marginal means for Healthy Controls and PD groups were, respectively, -0.44 and -0.29,[24] indicating that, although not statistically significant, individuals with PD had longer smile onset durations, even after controlling for smile intensity.

A*t*-test was also conducted to compare baseline log10-transformed AU12 Onset Duration data; Healthy Controls and individuals with PD did not significantly differ, $t(30) = -0.34$, $p = .74$.

Although the sample size was limited, Onset Duration was also examined separately for women (HC: $n = 6$; PD: $n = 9$) and men (HC: $n = 3$. PD: $n = 24$). ANCOVAs were performed with log10-transformed AU12 Onset Duration as the dependent variable, Group (HC and PD) as the between subjects factor, and AU12 Intensity as the covariate. There was no significant effect of Group for either men or women.*T*-tests performed for both men and women, revealed no significant differences between HC and PD groups.

---

[24] Values are negative due to log10-transformation of fractions. More positive values indicate longer onset duration.

Smiling behavior: impact of treatment upon frequency measures.

*Impact of treatment upon smile frequency: Post-Treatment (Time 2) versus Baseline (Time1).*Separate Kruskal-Wallis procedures were used to evaluate the impact of treatment group (Healthy Control, PD-Untreated, Artic, and LSVT) upon Duchenne Number, Non-Duchenne Number, Total Smile Number, and Duchenne Index; for each analysis, change-scores (T2-T1) were used as the dependent variable. See Table 20. There were no significant differences between the 4 treatment groups in change-scores for Duchenne Number, χ^2 (3) = 3.22, p = .36, Total Smile Number,χ^2 (3) = 4.66, p = .20, and Duchenne Index, χ^2 (3) = 4.47, p = .26). Non-Duchenne Number was not significant at the alpha level of .05 but did exhibit a trend, χ^2 (3) = 7.11, p = .07; LSVT had the highest mean change-score rank (34.53), followed by PD-Untreated (30.41), ARTIC(27.12), and Healthy Controls (18.27). Planned unadjusted pairwise comparisons of Non-Duchenne Number change-scores (T2-T1) revealed that LSVT did not significantly differ from ARTIC or PD-Untreated. On an exploratory basis, remaining pairwise comparisons were also examined. The mean rank of PD-Untreated change-scores (*Mdn* = 0.00, *M* = 0.41) was significantly greater than the mean rank of Healthy Control change-scores (*Mdn* = -1.00, *M* = -1.00), p = .05, and the mean rank of LSVT change-scores (*Mdn* = 1.00, *M* = 0.69) was significantly greater that Healthy Controls. There were no other statistically significant pairwise comparisons. SeeTable 20. Although gender effects may have contributed, in part, to observed findings, treatment analyses were not performed separately for each Gender due to limited sample sizes.

On an exploratory basis,Time 1 scores were directly compared to Time 2 scoresusing separate Wilcoxon signed-ranks tests for each dependent variable in each treatment group. See Table 21. For Duchenne Number, there was no statistically significant difference between pre-

and post-treatment scores for any of the treatment groups. For Non-Duchenne Number, a trend-level difference, $z = 1.78$, $p = .07$, was observed for the LSVT group, such that post-treatment scores were greater than pre-treatment scores; no other treatment groups exhibited significant differences; the same pattern was observed for Total Smile Number, also at a trend-level, $z = 1.67$, $p = .10$. For Duchenne Index, no treatment group exhibited a statistically significant change from pre- to post-treatment.

*Impact of treatment upon smile frequency: Follow-Up (Time 3) versus Baseline (Time 1).*Separate Kruskal-Wallis procedures were used to evaluate the relative impact—at Follow-Up, compared to Baseline—of each treatment group (Healthy Control, PD-Untreated, Artic, and LSVT) upon Duchenne Number, Non-Duchenne Number, Total Smile Number, and Duchenne Index. See Table 20. For each analysis, the change-score (T3-T1) was used as the dependent variable. There were no statistically significant differences between treatment groups for Duchenne Number, χ^2 (3) = 3.56, $p = .30$, Non-Duchenne Number, χ^2 (3) = 1.16, $p = .76$, Total Smile Number, χ^2 (3) = 2.25, $p = .52$, and Duchenne Index, χ^2 (3) = 3.77, $p = .29$.

Impact of treatment upon smile frequency: Follow-Up (Time 3) versus Post-Treatment (Time 2). Separate Kruskal-Wallis procedures were used to evaluate treatment group (Healthy Control, PD-Untreated, ARTIC, and LSVT) differences in change-scores, moving from post-treatment (Time 2) to 6-month follow-up (Time 3). For each of the frequency outcome measures, there were no group differences in Time 3 minus Time 2 change-scores: Duchenne Number, $\chi2(3, n = 53) = 2.03$, $p = .57$; Non-Duchenne Number, χ^2 (3, $n = 56$) = 2.49, $p = .48$; Total Smile Number, χ^2 (3, $n = 53$) = 2.55, $p = .47$; Duchenne Index, χ^2 (3, $n = 53$) = 2.00, $p = .57$.

Smiling behavior: impact of treatment upon intensity measures. To examine the

impact of treatment upon smile intensity, separate Kruskal-Wallis procedures were performed on each intensity outcome measures (AU6 Intensity, Duchenne AU12 Intensity, and Non-Duchenne AU12 Intensity),with treatment group as the independent factor. To assess group difference in the degree of change from baseline to post-treatment, the dependent variable was, for each outcome measure, Time 2 score minus Time 1 score. To assess group differences in the degree of change from baseline and post-treatment to follow-up, the dependent variables for each outcome measure, was, respectively, Time 3 Intensity minus Time 1 Intensity, and Time 3 Intensity minus Time 2 Intensity.

Impact of treatment upon smile intensity: Post-Treatment (Time 2) versus Baseline (Time1). For each intensity measure(AU6 Intensity, Duchenne AU12 Intensity, Non-Duchenne AU12 Intensity, and Overall AU 12 Intensity), separate Kruskal-Wallis tests were used to compare T2-T1 change-scores across treatment groups (Healthy Control, PD-Untreated, Artic, and LSVT). There were no statistically significant differences in T2-T1 change-scores between treatment groups for AU6 Intensity, χ^2 (3) = 3.56, p = .31, Duchenne AU12 Intensity, χ^2 (3) = 4.36, p = .23, Non-Duchenne AU12 Intensity, χ^2 (3) = 4.33, p = .23, and Overall AU12 Intensity, χ^2 (3) = 5.69, p = .13.

On an exploratory basis, pre- and post-treatment Intensity scores were directly compared using Wilcoxon signed-rank tests. For AU6 Intensity, a trend-level difference, z = -1.79, p = .07, was observed such that the PD-Untreated group had lower scores at post-treatment; no other treatment group exhibited a significant difference from pre- to post-treatment. For Duchenne AU12 Intensity, no treatment groups exhibited significant differences from pre- to post-treatment. For Non-Duchenne AU12 Intensity, the LSVT group exhibited a trend-level difference such that scores were higher after treatment, z = 1.71, p = .09; no other treatment

groups exhibited significant differences from pre- to post-treatment. Overall AU12 Intensity was also examined: in the LSVT group, post-treatment scores were statistically significantly greater than pre-treatment scores, $z = 2.51$, $p = .01$; there was a trend-level difference in the Healthy Control group, $z = 1.68$, $p = .09$, such that Time 2 scores were higher than Time 1 scores; no other treatment group exhibited significant differences in pre- and post-treatment scores.

Impact of treatment upon smile intensity: Follow-Up (Time 3) versus Baseline (Time 1). For each intensity measure(AU6 Intensity, Duchenne AU12 Intensity, Non-Duchenne AU12 Intensity, and Overall AU 12 Intensity), separate Kruskal-Wallis tests were used to compare T3-T1 change-scores across treatment group (Healthy Control, PD-Untreated, Artic, and LSVT). There were no statistically significant differences in T3-T1 change-scores between treatment groups for Non-Duchenne AU12 Intensity, χ^2 (3) = 3.87, $p = .28$, and Overall AU12 Intensity, χ^2 (3) = 2.21, $p = .53$.

For AU6 Intensity, there was a trend,χ^2 (3) = 7.11, $p = .07$, such that the mean change-score rank was highest for ARTIC (30.38), followed by LSVT (30.33), Healthy Controls (30.95), and PD-Untreated (18.88). Planned comparisons revealed that: LSVT change-scores ($M = 0.69$, $Mdn = 0.00$) were significantly greater, unadjusted $p = .03$, than PD-Untreated ($M = -0.66$, $Mdn = 0.00$); LSVT and ARTIC did not significantly differ, $p = .99$. As an exploratory measure, the remaining paired comparisons were also examined. The mean rank of ARTIC change-scores ($M = 0.43$, $Mdn = 0.63$) was greater than Untreated PD, $p = .04$; and the mean rank of Healthy Controls ($M = 0.67$ $Mdn = 0.40$) was greater than the mean rank of PD-Untreated.

For Duchenne AU12 Intensity, there was also a trend, χ^2 (3) = 7.05, $p = .07$, such that the mean change-score rank was highest for LSVT (33.10), followed by Healthy Controls (30.05),ARTIC (25.67), and PD-Untreated (20.38). Planned comparisons revealed that LSVT

T3-T1 change-scores (M = 1.11, Mdn = 0.00) were significantly, p = .01, greater than PD-Untreated (M = -0.25, Mdn = 0.00); LSVT and ARTIC (mean rank = 25.67) did not exhibit a statistically significant difference. The remaining pairwise comparisons were also examined. The mean rank of Healthy Controls (M = 0.62, Mdn = 0.50) was greater than PD-Untreated at a trend-level, p = .09; there were no other group differences.

Impact of treatment upon smile intensity: Follow-Up (Time 3) versus Baseline (Time 2). Change-scores were also used to assess Treatment group differences in intensity changes from Time 2 to Time 3. For each intensity measure, (AU6 Intensity, Duchenne AU12 Intensity, Non-Duchenne AU12 Intensity, and Overall Smile Intensity), separate Kruskal-Wallis tests were used to compare T3-T2 change-scores across Treatment groups (Healthy Controls, PD-Untreated, ARTIC, and LSVT). For each measure of intensity, there were no statistically significant differences between treatment groups: AU6 Intensity, $\chi2(3, n = 53) = 1.49$, p = .60; Duchenne AU12 Intensity, $\chi2(3, n = 53) = 1.86$, p = .60; Non-Duchenne AU12 Intensity, $\chi^2(3, n = 53) = .33$, p = .95; Overall AU12 Intensity, χ^2 $(3, n = 53) = 4.72$, p = .19.

Smiling behavior: impact of treatment upon temporal measures. To determine if the temporal characteristics of smile components were affected by treatment, a 2 (Time) by 4 (Treatment Group) ANOVA was performed with AU12 onset duration as the dependent measure. There were no statistically significant main effect of Time, $F(1, 25) = .08$, p = 0.78, partial η^2 = <.01, Treatment Group, $F(3, 25) = .47$, p = 0.7, partial η^2 = .05 nor was there a statistically significant interaction, Treatment Group*Time, $F(3, 25) = .46$, p = 0.72, partial η^2 = .05. On an exploratory basis, a Kruskal-Wallis was performed on the untransformed onset duration change-scores; change-scores did not significantly differ across treatment groups, p = 0.38. As there was no consistent correlation between AU12 Intensity and AU12 Onset Duration at different time

points, no steps were taken to control for AU12 Intensity.

Discussion

The human face is a vital channel for the expression of emotional life. Facial expressions complement, supplement, and, at times, substitute for spoken language. When facial expression falters, the accurate translation of subjective inner life into observable meaningful nonverbal signals is compromised. As a result, others may form inaccurate perceptions of one's personality, intelligence, or character, which, in turn, may negatively impact quality of life. Beyond the interpersonal significance of intact facial expressivity, there is evidence that facial expression might *directly* modulate one's emotional experience (see, e.g., Soussignan, 2002 for a recent discussion of the Facial Feedback Hypothesis). That our facial expressions are an important component of emotion communication, impression formation, and perhaps even one's own subjective emotional state, attests to the clinical significance of deficits in facial expressivity.

Reduced facial expressivity is common in Parkinson's disease (PD). Although the so-called "masked" presentation in PD is well-documented, our knowledge of causal mechanisms and potential treatments is limited. Using the Facial Action Coding System, this study comprised two main objectives. First, by measuring smile frequency, intensity, and onset duration, this study sought to extend our knowledge of baseline smile behavior differences between individuals with PD and healthy individuals. We also sought to evaluate treatment effects upon smiling behavior. Second, building upon prior investigations within the Emotion Lab at Queens College, this study sought to assess the long-term impact of treatment upon global measures of facial expressivity (Lability, Variability, and Complexity).

Baseline Differences between People with Parkinson's Disease and Healthy Controls

With respect to baseline differences in emotional expression, this study sought to replicate and extend the extant data on smile behavior in people with PD and healthy individuals. Three broad categories of smiling behavior were examined—frequency, intensity, and onset duration—with attention to the distinction between Duchenne and non-Duchenne smiles.

Number of smiles at baseline. In accord with previous studies, it was expected that, relative to controls, individuals with PD would exhibit reduced frequency of Duchenne smiling and a lesser amount of Duchenne smiles, relative to non-Duchenne smiles. As hypothesized, individuals with PD exhibited a lesser amount of Duchenne smiling, as compared to Healthy Controls. This finding is consistent with prior research using both FACS and impressionistic rating systems (Simons et al., 2003; Smith et al., 1996). In contrast to the hypothesized outcome, and in contrast to the findings of previous research demonstrating approximately equal overall smile number for PD and control groups (Pitcairn et al., 1990; Simons et al., 2003; Smith et al., 1996), the current study found that Total Smile Number was significantly greater in Healthy Controls as compared to individuals with PD (as was the number of non-Duchenne smiles). Similarly, contrary to hypotheses, the proportion of Duchenne to non-Duchenne smiles (as measured by the Duchenne Index) did not significantly differ between PD and control groups.

With respect to overall smile number, although hypothesized to not be statistically different between PD and control groups, previous research does generally show higher total smiling numbers for control participants. For example, although the Simons and colleagues study showed a pronounced discrepancy in terms of Duchenne smiling between controls (16 of 19 produced Duchenne smiles) and PD participants (4 of 15 participants produced Duchenne smiles), the level of total smiling also differed between groups, albeit less dramatically (73% of PD participants smiled in response to receiving a gift, in contrast to 84% of controls). Smith and

colleagues (1996) observed a similar pattern, such that, in a happy/spontaneous emotion elicitation condition, smiles with cheek raises were observed in 10/12 controls and only 4/12 PD participants; total smiles (with or without a cheek raise) were observed in 11/12 controls and 8/12 individuals with PD. Finally, Pitcairn and colleagues (1990) also observed significant differences in Duchenne smiling between individuals with ischemic heart disease (Cardiac controls) and individuals with PD, as well as absolute differences, albeit statistically insignificant in mean frequency (per minute) of total smiles (Cardiac controls [*N*=4] = 2.37; PD [*N*=4] = 1.81). Thus, obtaining statistically significant differences in terms of levels of mean overall smile frequency (HC = 3.09/min; PD = 1.38/min) in the current study appears to reflect a difference in degree (as opposed to direction), perhaps in combination with increased statistical power, rather than a conflict with the existing research.

Another possibility is that methodological differences in smile elicitation resulted in subtle but measurable differences in smiling behavior across studies. In the current study, smiles were elicited through participants' production of happy monologues. Other studies, however, employed a variety of other emotion elicitation techniques (e.g., semi-structured interview questions and description of the most recent vacation taken by the patient [Pitcairn et al., 1990], emotionally evocative film segments and requests to pose emotions [Smith et al., 1996], and spontaneous and posed reaction to odors, facial imitation, posed emotional expression, and facial activity measurement during conversation [Simons et al., 2003]).

Differing interpersonal relationships between subjects and experimenters might also account for differences in smile behavior. As noted by Halfacre et al. (2009), one possible explanation to account for prior findings of fewer Duchenne smiles coupled with approximately equal total numbers of smiles, is that PD participants are aware of their diminished spontaneous

smiling (i.e., Duchenne smiling) and attempt to compensate with voluntary (i.e., posed or non-spontaneous) facial expressions (Halfacre et al., 2009). This explanation is consistent with the finding that individuals with PD and without PD provide comparable ratings when asked to describe their own level of facial expressivity (Borod et al., 2008; Simons et al., 2004). Accordingly, it is possible that, in prior studies, individuals with PD felt more pressured by social emotion display conventions to deliberately produce non-Duchenne smiles to compensate for a relative dearth of spontaneously produced Duchenne smiles, whereas, in the current study, individuals with PD felt less compelled to produce compensatory non-Duchenne smiles. The extent to which different experimental procedures and interpersonal factors could have impacted the need to conform to social display conventions (e.g., deliberately smiling in order to present oneself as pleasant, polite, etc.) is unclear.

The equivalence of disease severity across studies was also considered. Overall, the average PD severity was similar across the current study and previous research. In the current study, at baseline, the average Hoehn and Yahr stage was 2.2. In the study by Simons and colleagues (2003), the average stage was 2.4,[25]PD severity was "mild to moderate" in the study by Pitcairn and colleagues (1990), and the overall average Hoehn and Yahr stage in the Smith and colleagues study (1996) was 2.5. Notably, when Pitcairn and colleagues separately analyzed the data for mild and moderate PD severity groups, they found that the proportion of smiles with a cheek raise (as compared to smiles without a cheek raise or no smiles at all) for the moderate PD group significantly differed from the control group ($p<.05$) whereas the mild PD severity group only marginally differed from the control group ($p=.07$).

[25] Calculated based on the information provided by Simons et al., 2003.

The possible impact of differing smile classification systems was also considered. In the current study, Duchenne and non-Duchenne smiles were explicitly classified as such based upon their FACS action units; specifically, Duchenne smiles were defined as the co-occurrence of action units 12 (AU12: zygomaticus major [lip corner puller]) and 6 (AU6: orbicularis oculi—pars orbitalis [cheek raiser]), whereas non-Duchenne smiles were defined as AU12 in the absence of AU6. Similarly, Simons and colleagues (2003) and Smith and colleagues (1990) identified "positive" smiles as the co-occurrence of AU 6 and AU12. Although Pitcairn and colleagues (1990) did require upturning of the lip corners and cheek raising with "crows' feet" (which would be commensurate with the FACS referents of AU6 and 12) for "happy" smile classifications, mouth opening and teeth exposure were also required. As the smile definitions employed by other researchers were either similar, or differed in such a way that would not result in more similar overall smiling between controls and PD participants, inter-study differences in smile classification are not thought to explain the current study's finding of significantly fewer overall smiles.

Finally, differing control populations may have contributed to differing smile patterns in PD relative to control groups. Although, like the present study, Smith and colleagues recruited healthy individuals for their control group, other researchers recruited medical controls (e.g., Pitcairn and colleagues used individuals with ischemic heart disease as a control group, and Simons and colleagues recruited healthy individuals, as well as individuals from medical support groups, as controls). It is possible that, with respect to smiling behavior, these control groups may not be equivalent to the healthy controls used in the current study. Specifically, it is possible that control populations consisting of individuals with chronic medical conditions

produce fewer smiles overall, thus increasing the likelihood of similar overall smiling frequency between individuals with PDs and (non-healthy) controls.

Intensity of smiles at baseline. In accord with previous findings (Smith et al., 1996), individuals with PD did exhibit less intense non-Duchenne smiles. Overall smile intensity was also less intense, that is, at a trend level, in the PD group. Contrary to expectations, individuals with PD did not significantly differ from healthy controls in either Duchenne AU12 intensity (AU12s co-occurring with AU6) or in AU6 Intensity (AU6 co-occurring with or without AU12). The absence of group differences in Duchenne AU12 intensity is especially surprising in light of the common view that, in PD, spontaneous expression (which would typically include Duchenne smiles) is more adversely impacted than posed expression. Notably, although not reaching statistical significance, the mean intensity values for both AU6 and Duchenne AU12 were lower in the PD group than in the healthy control group.

Temporal features of smiling in PD and healthy controls. In accordance with previous literature (Bowers, Miller, Bosch, et al., 2006), it was predicted that PD participants would demonstrate bradykinetic facial expression. Contrary to expectation, PD and healthy control groups did not significantly differ in smile onset duration. Because the onset duration data set within the current study was relatively small (due to the limited number of smiles whose onset phase fell entirely within the coded section of monologues; see Results section above), the current study may have lacked sufficient power to reveal subtle differences in onset duration (0.49 seconds, on average, for healthy controls; 0.56 seconds for individuals with PD).

With respect to potential cohort differences between the studies, disease severity was similar between the Bowers et al. (2006) study (average Hoehn-Yahr stage was 2.8) and the current study(average Hoehn-Yahr stage was 2.2). Methodological differences, however, might

explain the different results between the two studies. In the present research, onset duration was defined as the amount time elapsed, from smile onset to the first peak of emotional expression (i.e., if a smile appeared to undergo distinct successive intensity increases, the first peak was used to define the end of the onset phase). Bowers and colleagues, however, defined the onset phase as the time to reach peak expressivity relative to the sound of a buzzer, which signaled participants to generate a facial expression. Thus, whereas the period of time measured by Bowers and colleagues might have reflected some amount of reaction time delay (e.g., resulting from perceptual processing, cognitive processing, or motor planning), the current study measured only motor execution duration (i.e., excluding pre-movement processing).

Additionally, Bowers and colleagues measured posed facial expressions whereas the current study measured naturalistic expressions. Accordingly, although the current study may be more generalizable to everyday contexts—insofar as naturalistic expression typically incorporates a combination of spontaneous expression as well as deliberately modulated and/or masking of facial expression—Bowers and colleagues may have achieved a greater degree of experimental controlby restricting analysis to posed expression. However, due to the relative infrequency of Duchenne smiles in the current study for which onset duration could actually be calculated (and thus, the predominance of non-Duchenne smiles, which are relatively more likely to be posed), the dataset analyzed in the present study may be largely comparable to that of Bowers and colleagues.

The methodological differences between the current study and that of Bowers and colleagues is salient in light of the pathophysiological features of bradykinesia in PD; namely, as argued in a literature review by Berardelli and colleagues (Berardelli, Rothwell, Thompson, &Hallert, 2001), bradykinesia “results from a failure of basal ganglia output to reinforce the

cortical mechanisms that prepare and execute the commands to move…This leads to particular difficulty with self-paced movements, prolonged reaction times, and abnormal pre-movement EEG activity" (Berardelli et al., 2001, p. 2131). Thus, the extent to which previously reported dynamic facial expressivity deficits in PD might reflect a combination of slowed movement (bradykinesia), versus poverty of movement and/or slowness to initiate movement (akinesia) is unclear.

Finally, because the onset duration data set within the current study was relatively small (due to the relatively small number of smiles whose onset phase fell entirely within the coded section of monologues; see Results section above), it was not possible to separately analyze Duchenne and non-Duchenne smiles. If separate analyses were possible, consistent with the Bowers and colleagues findings, as well the more general fact of slowed movement in PD, it would have been expected that PD participants would be bradykinetic for both Duchenne and non-Duchenne smiles. If, however, the temporal dynamics of Duchenne and non-Duchenne smiles were differentially impacted, this would be informative with respect to the mechanism by which PD affects facial expressivity.

Treatment Effects

Beyond describing baseline facial expressivity differences between individuals with PD and healthy controls, this study explored treatment effects of LSVT LOUD and an articulation-focused control intervention (ARTIC) in PD relative to healthy controls and to individuals with PD who received no treatment. As with the examination of baseline characteristics, treatment effects were explored in relation to their impact upon smiling behavior, including smile frequency, intensity, and duration, as well as an examination of Duchenne versus non-Duchenne smiles. Furthermore, extending the pre-/post-treatment findings regarding global measures of

facial expressivity—lability, variability, and complexity (Dumer et al., 2010)were examined at a 6-month follow-up. Although 6-month follow-up data were collected for smile outcome measures, we focused on baseline versus post-treatment findings in view of possible systematic differences in 6-month follow-up coding vis-à-vis baseline and post-treatment coding. With respect to inter-rater reliability, general challenges (i.e., those common to FACS coding) and obstacles unique to this study are discussed below following a summary of treatment findings.

Smiling behavior in treated and untreated Parkinson's disease.

Number of smiles after treatment. It was predicted that LSVT LOUD, and, to a lesser extent, ARTIC, would result in an increased number of smiles. Furthermore, it was predicted that LSVT would increase the Duchenne Index (i.e., number of Duchenne smiles relative to non-Duchenne smiles) whereas ARTIC would not. If this differential effect of treatment uponDuchenne versus non-Duchenne smiling was to be observed, it would have demonstrated that the effect of LSVT LOUD is *not* restricted to speech-specific AUs, but insteadaffects both speech *and* non-speech AUs (e.g., AU6, cheek-raiser, an upper-face AU). Such a finding would have supported the hypothesis that LSVT, through targeting effort and vocal loudness as opposed to strictly orofacial musculature, activates motor pathways shared by both speech and non-speech output (see Fox et al., 2006 for discussion of potential mechanisms of action).

To assess these hypotheses, for each variable of interest, change-scores were examined across treatment groups, and scores at Times 1 and 2 were directly compared. In contrast to hypothesized treatment effects, there were no treatment group differences in either overall smile number change-scores or Duchenne Index change-scores. The only frequency variable demonstrating between group differences in change-scores was Non-Duchenne Number. As revealed by post-hoc comparisons, consistent with hypotheses, LSVT change-scores were

significantly greater than those of Healthy Controls; PD-Untreated change-scores were also greater than those of Healthy Controls. Inspection of group means and medians revealed that, whereas both the LSVT and PD-Untreated groups did exhibit increased scores (i.e., positive change-scores), Healthy Controls exhibited lower scores at Time 2 relative to Time 1 (i.e., negative change-scores).

With respect to direct within-treatment group comparisons of Time 1 and Time 2 scores, only the LSVT group demonstrated trend-level increases. Within the LSVT group, there were trend-level increases, from baseline to post-treatment, in both non-Duchenne number and total smile number. Although the mean score for Duchenne number increased from baseline to post-treatment, this increase was not statistically significant. LSVT also did not exhibit the hypothesized increase in Duchenne Index. Thus, the current findings suggest that LSVT increases overall smile frequency, largely through an increase in non-Duchenne smiles. There was, however, no support for the hypothesis that LSVT would normalize the proportion of non-Duchenne to Duchenne smiles, and there were inconclusive findings with respect to the potential of LSVT to impact muscle activity remote from those directly involved in speech (as evidenced by a mean increase in Duchenne Number that failed to reach statistical significance).

Intensity of smiles after treatment. It was predicted that LSVT LOUD and, to a lesser extent, ARTIC, would result in increased smile intensity. As described above, it was hypothesized that the effect of LSVT LOUD would notbe restricted to speech-specific and adjacent AUs but would, instead, affect both speech *and* non-speech facial musculature. This hypothesis would be supported if LSVT LOUD resulted in increased intensity of AU 6 (cheek raiser; a marker of emotional expression that is not required for speech) at Time 2, compared to Time 1.

As predicted, from Time 1 (baseline) to Time 2 (post-treatment), LSVT resulted in a statistically significant overall increase in smile intensity (as was the case for the HC group), as measured by increases in AU 12 intensity, regardless of absence or co-occurrence of AU6. Also consistent with predictions, non-Duchenne intensity increased in the LSVT group at a trend-level. Although the mean intensity of both AU6 and Duchenne smile intensity (AU 12 in the presence of AU6) did increase from baseline to post-treatment, contrary to the hypothesized outcome, neither exhibited statistically significant or trend-level increases at Time 2. ARTIC, hypothesized to exhibit more mild increases in intensity, did not exhibit statistically significant or trend-level intensity increases.

In a separate set of analyses, change-scores were examined across each of the four treatment groups. For Time 1 versus Time 2 and for Time 2 versus Time 3 change-scores, there were no treatment group differences in intensity. Interestingly, for Time 1 to Time 3 change-scores, group differences were obtained at a trend-level for AU 6 and Duchenne AU 12 intensity, such that LSVT increases were greater than untreated PD scores. Healthy control change-scores were also significantly greater than the untreated PD group for these measures of intensity; LSVT did not significantly differ from ARTIC, and no other group differences were obtained.

The impact of treatment on temporal features of facial expression. Finally, this study explored the impact of treatment upon smile onset duration. All else being equal, to the extent that treatment increases expressive intensity, one would expect a corresponding increase in onset duration; in other words, if intensity increased and onset speed remained constant, and there were no other variables affecting facial expression, it would necessarily take a longer amount of time to reach peak intensity. On the other hand, through exercise effects upon facial muscles or more central treatment effects impacting basal-ganglia related pre-motor (e.g., movement "scaling,;"

Alexander, Crutcher, & DeLong, 1990), intervention could also result in more efficient (i.e., more rapidly executed) facial expression. In the current study, however, there was no significant treatment effect upon onset duration, even when controlling for expressive intensity.

As noted above, due to limitations in sample size, onset duration could not be separately analyzed in terms of Duchenne and non-Duchenne smiles. Had the sample permitted a separate analysis, an examination of Duchenne versus non-Duchenne smile dynamics would have been informative in light of research demonstrating differences in the temporal features of posed and spontaneous facial expressions, which roughly correspond to non-Duchenne and Duchenne smiling. For example, slower smile onsets have been associated with spontaneous smiles (e.g., Schmidt et al., 2003) and positive impression formation (Krumhuber et al., 2007). With respect to the overall analysis, however, because treatment groups did not significantly differ in the proportion of Duchenne to non-Duchenne smiles, differing temporal characteristics of Duchenne versus non-Duchenne smiles are likely not a significant confound, and one can thus cautiously interpret the onset duration results and conclude that overall smile onset duration was not substantially impacted by treatment.

For future investigations of treatments targeting facial expressivity in PD, it is interesting to note that in PD, facial expression can be impacted by bradykinesia and can lead to negative impression formation, whereas, as noted above, in the general population, slower smile onsets (approximately 0.5 to 0.75 seconds) have been associated with spontaneous smiles (e.g., Schmidt et al., 2003) and positive impression formation (Krumhuber et al., 2007). Thus, to the extent that the bradykinetic facial expressivity in PD can be effectively targeted, it would be important to evaluate the impact such a change would have on interpersonal impression formation (i.e., in terms of socially desirable characteristics).

Global measures of facial expressivity at 6-Month follow-up. As evidenced by previous research in the Emotion Lab at Queens College (Dumer, 2011), select global measures of facial expressivity (i.e., Lability and Variability) were higher—at either statistically significant or trend-levels—within the LSVT group as compared to the PD-Untreated group, ARTIC, and Healthy Controls. In addition to the previously described findings regarding smiling behavior, the current study assessed the long-term (6-month) efficacy of treatmentupon global measures of facial expressivity in PD.

In advance of Dumer's (2011) findings, it was hypothesized that at Time 3 (i.e., 6-month follow-up), ARTIC and LSVT treatment conditions would exhibit statistically significant increases in facial expressivity compared to Time 1 (pre-treatment). Furthermore, it was predicted that the effect of LSVT LOUD would be significantly greater than the effect of ARTIC. Finally, because there was no active treatment between Time 2 (post-treatment) and Time 3 (follow-up), it was predicted that treatment effects (i.e., increased facial expressivity) of both ARTIC and LSVT LOUD would diminish such that the values of the global facial expressivity measures at Time 3 would be significantly less than the values obtained at Time 2. Following Dumer's findings, it was hypothesized that same pattern of treatment effects observed at Time 2 would persist, but be diminished, at 6-month follow-up (Time 3).

Contrary to hypotheses, our findings suggested that none of the four treatment groups (Healthy Controls, PD-Untreated, Artic, or LSVT) significantly differed with respect to degree of change at Time 3 relative to Times 1 and 2. This was also the case for each global expressivity variable (Lability, Variability, and Complexity) and for each change-score examined (i.e., Time 3 minus Time, and Time 3 minus Time 2). Thus, we found no evidence of group

differences in global measures of facial expressivity at the 6-month follow-up, as evidenced by the absence of statistically significant differences in change-scores. These results contrast with hypothesized outcome, namely, that LSVT and, to a lesser degree, ARTIC, would exhibit greater treatment effects at the 6-month follow-up, as compared to Healthy Controls and the untreated Parkinson group.

Study Limitations.

Analyzing follow-up scores in relation to baseline and post-treatment. The current analyses were targeted at detecting group differences in degree of change over time. Due to potential systematic differences in coder scoring which could complicate direct comparison of absolute levels of facial expressivity at Time 3 against either absolute levels of facial expressivity at Time 1 or Time 2, the assessment of facial expressivity at 6-month follow-up was primarily examined using change-scores (i.e., comparing degree of change across groups), and was not followed up with statistical tests directly comparing Time 3 scores to Time 1 or Time 2 scores[26]. Notably, null findings for Time 3 change-score analyses across groups does not preclude the possibility that, within one or more treatment groups, facial expressivity at 6-month follow-up outcomes was meaningfully different, relative to pre- and/or post-treatment. In fact, for some variables, the possibility of increased facial expressivity at Time 3 would be, prima facie, consistent with the data (e.g., Lability and Variability scores were consistently higher at Time 3 relative to Time 1 and 2). However, an actual increase in facial expressivity from Time 1 and 2 to Time 3 is judged to be unlikely due to 1.) lack of a theoretically driven explanation to account for increased facial expressivity at Time 3, and 2.) potential systematic inter-coder

[26] Change-score was selected in light of the failure of the dependent measures to meet parametric assumptions, but confers the benefit of permitting meaningful analysis of group differences despite potential systematic coding differences at Time 3 as compared to Times 1 and 2.

differences accounting for higher Time 3 scores (coded by D.M.) that might better account for apparent Time 3 increases.

With respect to the trajectory of facial expressivity over time, two hypotheses appear to be more obvious and parsimonious. On the one hand, Time 3 scores might remain unchanged relative to Time 2, thus demonstrating a stable treatment effect (or lack of treatment effect, if no change was observed at Time 2 relative to Time 1). A second hypothesis is that Time 3 scores would demonstrate persistent albeit, diminished treatment effects, resulting from cessation of treatment and/or progression of PD. For variables demonstrating higher Time 3 scores relative to Time 1 and 2 scores (e.g., Lability and Variability), there are at least two factors that might account for this increased expressivity over time. First, it is at least conceivable that, at Time 3, subjects were more expressive due to a cohort effect, (e.g., increasingly comfortable relationship between the subject and the individual prompting the elicitation of Happy monologues). Second, of more theoretical and clinical interest, is the possibility that, from post-treatment to 6-month follow-up, subjects—deliberately or subconsciously—practiced the interventions delivered in the LSVT and or ARTIC conditions, even after the completion of formal treatment. However, for variables demonstrating higher Time 3 scores, inter-coder differences might better account for higher Time 3 scores.

Challenges to inter-coder reliability. Inter-coder reliability is particularly important when comparing data sets coded by two different individuals—as is the case with Time 3 scores (FACS coded by D.M.) as compared to Time 1 or 2 scores (coded by A.D.). As global facial expressivity at the 6-month follow-up was a primary target of investigation in the current study, factors affecting inter-coder reliability are discussed below, with an emphasis on analysis of treatment effects upon global facial expressivity measures.

With respect to inter-coder reliability, although Time 3 intra-class correlation coefficients were relatively high for Lability and Variability, they were lower for Complexity, indicating a less consistent linear relationship in coders' ranking of Complexity scores. Although intra-class correlations provide an important metric of agreement, it is crucial to examine the data for potential systematic coding differences not revealed by the intra-class correlation statistics. For example, for both Lability and Variability, which demonstrated relatively high intra-class correlation coefficients, there appeared to be evidence of systematic coding differences. Specifically coder D.M., who coded the Time 3 data, tended to generate higher scores (D.M.'s average scores, for the reliability subset at Time 3, were 1.6 units higher for Lability, 1.9 units higher for Variability, and 0.19 units higher for Complexity). In light of the possibility that higher Time 3 scores could be accounted for systematic coder differences, qualitative or statistical analyses relying upon direct comparison of Time 3 scores relative to baseline or post-treatment scores (e.g., a within-group analysis of Time 2 scores to Time 3 scores) may lead to misleading conclusions. Notably, although potential time-dependent differences in coding limit the ability to make meaningful direct comparisons of Time 3 scores relative to Time 2 or Time 1 scores, the comparison of change-scores across treatment groups did permit a statistical approach to answering the question "Does one treatment group differ from another in the *degree* of change between two time points?" Having established that there likely were important inter-coder differences, it may be helpful to future investigators to appreciate potential sources of both systematic inter-coder differences as well as more general challenges to establishing and maintaining inter-coder reliability.

Training differences. It is notable that although both of the FACS coders in this study (D.M. and A.D.) were successfully trained to criteria on FACS, each coder learned FACS using a

different mode of training. Coder A.D. trained with a live cohort during an intensive in-person FACS training seminar. Coder D.M. followed a self-study program using the FACS manual. Although the coders later co-coded stimuli to establish inter-rater reliability, it is possible that some systematic differences in coding behavior were acquired as a result of different training methods.

Methodological challenges. Although FACS is designed to capture dynamic facial movement, it is worth noting that the current study necessarily involved suboptimal coding conditions. Favorable FACS coding conditions might include, for example, the use of standardized emotionally neutral photos to be compared against still-photos, or brief video-data documenting facial expressions originating from a neutral expression. The current study, however, involved the coding of naturalistic facial expression, which contains facial events originating from neutral and non-neutral states, produced during speech. Although some speech movements are clearly identifiable as such and could easily be screened from analysis, speech can, at times, involve action units that can be interpreted as emotionally laden (and thus codable) action units.

Whereas some facial actions might be readily identified, other situations may require judgment, reference to coding rules, and comparison of the to-be-coded stimuli against emotionally neutral frames. In the course of coding a monologue, a FACS coder might frequently refer to neutral frames in a number of situations. The exact relation of the neutral frame to the stimulus could potentially affect a number of aspects of coding such asdistinguishing between low-level AU activations from baseline facial morphology(e.g., baseline brow structure versus slight brow raise, naturally upturned lips versus slight smile, etc.), determining the exact timing of onsets and offsets, and distinguishing among similar AUs.

With respect to neutral frame selection, the neutral frames used in Time 1 and Time 2 coding were not documented. Accordingly, when coder D.M. later coded Time 3, he needed to independently identify neutral frames. Frequently, the choice of a neutral frame was obvious. At other times, however, the "correct" neutral frame was either not obvious, or an optimal neutral frame was not available (i.e., during the recorded monologue, the subject may have constantly produced a particular AU or combination of AUs). Potentially, even subtle differences in choice of neutral reference points could affect 1.) inter-coder reliability within the given coding subset and 2.) consistency of coding at Time 3, as compared to Time 1 and 2.

Still, there are other factors that may have impacted inter-coder reliability and/or actual coding differences at Time 3, relative to Times 1 and 2.With respect to inter-coder reliability, it is possible that inter-coder reliability was impacted by subtle perceptual biases arising from differing experimental hypotheses, such as, for example, coder D.M. was more sensitive to smile events and coder A.D. was more attuned to global facial expressivity. Concerning potential systematic differences between Time 3 as compared to Times 1 and 2, coding may have been impacted by subtle variations in lighting or camera angle, subject mood, or efficacy of emotion elicitation.

Gender. Another limitation of this study wasgender imbalance within the Healthy Control group (which contained more women) and PD group (which contained more men). This gender imbalance is important in light of the documented gender differences in emotional expressivity, such that women tend to be more emotionally expressive than men (e.g., Borod & Madigan, 2000; Grunwald et al., 1999; LaFrance et al., 2003). With respect to the current study, the implication of the gender imbalance in combination with possible gender differences in facial emotional expressivity, is that overall group findings (HCs versus PDs, regardless of gender)

could represent an artifact such that apparently higher Healthy Control scores, relative PD scores, might be better accounted for by the larger number of women (who, tend to be more emotionally expressive) in the HC group and the higher number of men (likely less emotionally expressive, on average) in the PD group.

As reported in the Results section, to better understand the impact of gender within the current data set, each facial expressivity variable was examined at baseline,to determine if scores differed as a function of gender; this analysis was conducted separately for the HC and PD groups. A number of variables exhibited gender differences, especially within the PD group. When gender differences were found, women exhibited higher levels of expressivity as compared to men.

To further explore the potential gender confound in comparing baseline HC and PD groups, gender-specific analyses were conducted, in addition the overall group analysis. When the baseline data were analyzed separately for men and women, a number of overall (i.e., HC versus PD) differences were no longer present. This suggests that gender is an important consideration but, notably, when outcome measures were analyzed separately for men and women, small sample sizes may have also contributed to gender-specific null findings.This pattern of findings underscores the necessityto exercise caution when assessing the generalizability of emotional expression findings—related to PD or any other clinical population—obtained from samples withan unequal balance of men and women, or from samples limited exclusively to one gender. Regarding studies of smile behavior in PD, a gender imbalance occurred for the current study (PD: 12 women, 33 men; Healthy controls: 7 women, 4 men),and Smith et al., 1996 (PD: 4 women, 8 men; Controls: 6 women, 6 men); the samplestudied by Pitcairn and colleagues(1990)was limited to men.The gender imbalance

sometimesseen in PD studies mayreflect a possibly higher risk of developing PD in men, as compared to women (e.g., Wooten, Currie, Bovbjerg, Lee, &Patrie, 2004).

Summary and Future Directions

To better understand the nature of facial expression deficits in Parkinson's disease, a two-component study was conducted. First, the long-term (6-month) efficacy of LSVT LOUD was compared to a second intervention (ARTIC), which targets articulation, in treating facial expressivity changes in PD. Global measures of facial expressivity were used to study 6-month follow-up data and build upon the pre-/post- findings of Dumer (2011). Second, smile behavior was examined at baseline and as a function of treatment condition and time. Smile frequency, intensity, and onset duration data were examined. Duchenne smiles (AU12, co-occurring with AU6), commonly thought to reflect spontaneous or "felt" emotion, were distinguished from non-Duchenne smiles (AU12, not in the presence of AU6). Facial expressions were coded using the Facial Action Coding System (FACS) developed by Ekman and Friesen (1978).

Summary of baseline smile behavior differences between individuals with PD and healthy controls. At baseline, in accord with previous findings, Healthy Controls generally exhibited higher levels of facial expressivity as compared to individuals with PD.With respect to measures of smile frequency, Healthy Controls did, as expected, exhibit moreDuchenne smiling. Healthy Controls also exhibited more non-Duchenne smiling and, unexpectedly, did not differ from individuals with PD in terms of the ratio of Duchenne to non-Duchenne smiling. With respect to smile intensity, Healthy Controls exhibited greater overall smile intensity but, contrary to expectations, did not significantly differ from individuals with PD in Duchenne smile intensity. Finally, groups did not significantly differ in terms of onset duration.

Because smile frequency, intensity, and onset duration were examined, the results of the current study can be described, respectively, in terms of akinesia (poverty of movement), hypokinesia (reduction of movement intensity), and bradykinesia (slowed movement). As a trend was observed, such that the number of both Duchenne and non-Duchenne smiles was lower in the PD group, PD participants appeared to demonstrateakinesiafor both types of smiles. As non-Duchenne smiles were significantly less intense in the PD group, hypokinesiaappeared to impactnon-Duchenne smiles, but not necessarily Duchenne smiles (which did not significantly differ between groups, but which were lower in terms of absolute value in the PD group). Finally, as there were no group differences in onset duration, there was no evidence of bradykinesia, as assessed by the current measurement paradigm. Notably, the onset duration measurement paradigm used in this studymeasured the amount of time elapsed from smile onset to the first peak of expressive intensity (i.e., a design thatshould be minimally impacted by deficits in reaction time or pre-execution motor planning processing); studies employing different temporal measurement paradigms (e.g., Bowers, Miller, Bosch, et al., 2006) may produce different findings. Additionally, the absence of group differences in onset duration might be attributable to a reduced sample size, which prevented a separate analysis of Duchenne and non-Duchenne smiles.Finally, as previously noted, the baseline findings between individuals with PD and healthy controlsshouldbe interpreted with caution in light of agender imbalance between the PD and healthy control groups, in combination with potential gender differences in facial expressivity.

In summary, considering the overall frequency, intensity, and onset duration data, thecurrent study further supports previous observations documenting a generalreduction of facial expressivity in PD.In light of prior research regarding spontaneous versus posed emotional

expression in PD and because Duchenne smiles are commonly thought to reflect spontaneous (or "felt" emotion) and non-Duchenne smilingare thought to reflect intentionally modulated (and,perhaps, posed emotion), a greater deficit was expected to be seen in Duchenne smiling. However, asthe current study was not designed to be a controlled study of spontaneous versus posed emotion, these findings do not necessarily challenge previous reports. The current results do, however, highlight the possibility that any disproportionate impact upon spontaneous relative to posed expression might not be evident in naturalistic paradigms, such as those employed in the current research.

Summary of treatment effects. Following treatment, several smile behavior variables exhibited statistically significant or trend-level increases. All increases were limited to the LSVT group with the exception of an overall increase of AU12 intensity within the Healthy Control Group. Analysis of change-scores, however, generally found no differences between treatment groups in terms of degree of change over time. Thus, the current results provide limited support for hypothesis that LSVT would demonstrate greater efficacy, as compared to other treatment groups, in increasing facial expressivity. However, because Duchenne number and Duchenne index did not significantly change in PD groups, there was no evidence that LSVT, or any other treatment, normalizes the overall number or ratio of Duchenne to non-Duchenne smiling. Furthermore, there was no evidence that LSVT, or any other treatment, impacts smile-relevant AUs beyond the orofacial region. In other words, the hypothesis that LSVT—through effortful targeting of the vocal loudness parameter, as opposed to the isolated targeting of articulation—would exert treatment effects outside of the orofacial region was not supported; similarly, there was no evidence that LSVT, relative to an articulation-based treatment, exerts a qualitatively different impact on facial expressivity.

At 6-month follow-up, global measures of facial expressivity did not significantly differ across treatment groups. Although the LSVT group increased in some measures of smile behavior, LSVT did not generally differ from other treatment conditions in degree of treatment impact over time, as assessed by a nonparametric analysis of change-scores. Accordingly, for global measures of facial expressivity (e.g., Lability and Variability), there is evidence that despite group differences from baseline to post-treatment, treatment group differences might diminish by the time of six-month follow-up. These results should be cautiously interpreted in light of inter-coder reliability challenges but should, nonetheless, be interpreted as preliminary evidence that facial expressivity benefits obtained immediately after treatment may diminish over time.

Future directions. In planning future studies, researchers should consider a number of factors. Anticipating potential gender effects, investigators should take steps to ensure gender balance across treatment groups and between the clinical and control samples. Samples sizes should be large enough to detect subtle yet potentially clinically meaningful treatment effects and to permit gender-specific analysis of findings.Additionally, researchers should be aware that when investigating smile behavior, or other non-aggregate measure of facial expressivity, the dependent measure might be low-frequency in its occurrence, thus affecting the distribution of the outcome data (i.e., resulting in a failure to meet parametric assumptions) and, in turn, constraining or complicating the appropriate statistical approach to data analysis.

Future investigations of expressivity deficits in PD can not only help to elucidate the neural mechanisms accounting for such deficits but can also be used to educate patients, their friends and families, and clinicians about the nature of facial expressivity deficits in PD. With respect to future studies of LSVT—or other treatments for which facial expression is not the sole

target behavior—it would be valuable to determine the extent to which facial expression increases are correlated with successful voice and speech outcome measures. Such findings would not only facilitate prediction of facial expression outcomes but might provide outcome markers to inform continuation or cessation of treatment.Also, although the focus of the current study was to evaluate collateral facial expressivity gains derived for a treatment targeting voice and speech deficits, investigators should further explore treatments that directly target facial expression deficits in Parkinson's disease. In light of the potential facial expressivity improvements following behavioral parameter targeting (i.e., LSVT), facial physiotherapy (Katsikitis and Pilowsky, 1996), and even music therapy (Elefant, Lotan, Baker, &Skeie, 2012), it would be important to identity which interventions yield clinically significant changes in facial expression. Furthermore, if different interventions yield qualitatively different treatment effects, it would be important to establish the feasibility of combining treatments or integrating select facets of treatment protocols. Lastly, researchers should consider whether treatment modifications might increase treatment efficacy (e.g., an LSVT-like method directly targeting intensity of facial expression).

Although clinicians may reasonably expect some facial expression increases following LSVT and other treatments, care should be taken to disclose to patients that an increase in facial expressivity may not translate into a full normalization of facial expression. Finally, clinicians should be aware that treatment effects observed immediately following an intervention may dissipate over time, thus arguing for the incorporation of clinician- and/or client administered "booster sessions" to maintain treatment effects over prolonged periods of time.

Table 1

Number of Subjects in each Treatment Condition at Times 1(Baseline), 2 (Post-Treatment), and 3 (6-Month Follow-Up)

	Time 1	Time 2	Time 3
HC	11 (7)	11 (7)	10 (6)
Untreated PD	17 (4)	17 (4)	16 (4)
Artic	12 (4)	12 (4)	12 (4)
LSVT	16 (4)	16 (4)	15 (4)

Note. Number of women in each group is provided in parentheses.

Table 2

Onset Duration N: Number of Subjects in Each Treatment Condition, in Each Time

	Time 1	Time 3
	Total *N* (# of Women)	Total *N* (# of Women)
PD (All)	23 (6)	22 (6)
HC	9 (6)	8 (5)
Untreated PD	7 (2)	8 (3)
Artic	8 (1)	6 (0)
LSVT	8 (3)	8 (3)

Table 3

Participant Demographics and Relevant Medical Variables

Variable		Treatment Group					
		ARTIC	LSVT	Untreated	Control	Total	*p*-value[3]
Sample Size		12	16	17	11	56	
Age[1]		69.3 (10.3)	68.5 (6.7)	65.7 (8.9)	61.8 (8.6)	66.5 (8.8)	.15
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)	.61
Hoehn-Yahr Stage[2]		2.2 (0.7)	2.2 (0.5)	2.1 (0.6)	N/A	2.2 (0.6)	.77
Years since diagnosis[2]		5.1 (4.0)	5.9 (7.1)	6.6 (5.9)	N/A	6.0 (5.8)	.70
# of PD medications[2]		1.7 (1.3)	2.0 (0.9)	2.3 (1.1)	N/A	2.0 (1.1)	

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)	<.01 HC <PD groups

[1] For the 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).
[3] For select variables, p-values are presented for analyses of treatment group differences. Other demographic characteristics are discussed within the text.

Table 4

Aim 1 Inter-rater Reliability

	Times 1 and 2[4]	Times 3
Overall percent agreement[1]	81%	66%
FACS Agreement Index[2]	.71	.40
Lability[3]	.80	.80
Variability[3]	.75	.81
Complexity[3]	.60	.63

Note.
[1]Percentage of agreement on occurrence of an event.
[2]Calculated as described in Ekman, Friesen, and Hager (2002a); FACS Agreement Index = ((# AUs on which coder AD and DM agreed) X 2)/(Total Number of AUs scored by both coders),
[3]Intra-class correlation coefficients (ICCs) of ranked transformed scores: ICCs for Time 3 calculated as part of present investigation. Single-measures coefficients are reported here as per Hallgren (2012); consistency definition values are reported. ICCs for Times 1 and 2 were originally reported in Dumer (2011).
[4]Aim 1 reliability at Times 1 and 2, as reported in Dumer (2011).

Table 5

Aim 2 Inter-rater Reliability

	Times 1 and 2	Times 3
Percentage Agreement by Smile Type[1]		
Non-Duchenne	.46, .73	.32, .66
Duchenne	.74, .74	.67, .75
Total Smile	.57, .79	.44, .74
Onset Duration (First Peak) Agreement[2]		
0.50 Second Tolerance Window[3]	.78	-
0.25 Second Tolerance Window[3]	.58	-
Intensity[4]		
AU6	.68*	*
AU12	.95*	*

Note.
[1]Percentage of AUs upon which coders agreed; Calculated as (#Agreements)/(#Agreements + #Disagreements). The first percentage requires coders to have identified the start time of a given AU within one second of the other coder; the second percentage defines agreement as any degree of same-AU overlap.
[2] In seconds, for agreeing smiles. Includes raw values, thus net average of +0.5 and -0.5 would be 0."First Peak" = First observable intensity peak.
[3] Agreement index = # Agreements/(#Agreements + #Disagreements). "Agreement" is defined as both coders coding within +/- intensity level (i.e., B vs. C --> agree; B vs. D does not agree.).
[4]For intensity, agreement was calculated across Times 1, 2, and 3.

Table 6

Normality, Homogeneity of Variance, and Outlier Summary for Global Expressivity Measures: Baseline Comparison of Healthy Control and Parkinson's disease groups

Variable (Time 1)		Distribution Normal in both Groups? ("3" if yes; blank if no.)	Homogeneity of Variance?	Outliers Absent?
Lability	Raw			
	Square root	3	3	3
	Log10	3	3	3
	Inverse			
Variability	Raw			
	Square root	3	3	PD: 64, 65@ T2; 5 @T3
	Log10	HC:.22; PD .03	3	PD: 64, 65@ T2, 58@T3
	Inverse			
Complexity	Raw			
	Square root			
	Log10			
	Inverse			

Note. First, normality was assessed using the Shapiro-Wilk test. If the distribution was normal in both the HC and PD group, homogeneity of variance was assessed with the Levene test. If homogeneity of variance was present, outliers were examined.

For Variability, neither square root nor log10 transformation fully met parametric assumptions. To understand the impact of transformation on score distributions, Group and Shapiro-Wilk test p-values are provided in the log10 row of the "Distribution Normal. . ." column. Additionally, Group and Time and subject number are indicated in the square root and log10 rows of the "Outliers Absent" column.

Table 7

Normality, Homogeneity of Variance, and Outlier Summary for Smile Behavior Measures: Baseline Comparison of Healthy Control and Parkinson's disease groups

Variable (Time 1)		Distribution Normal in both Groups? ("3" if yes; blank if no.)	Homogeneity of Variance?	Outliers Absent?
FREQUENCY		No normal distributions		
Duchenne #	Raw			
	Square root			
	Log10			
	Inverse			
Non-Duchenne #	Raw			
	Square root			
	Log10			
	Inverse			
Total Smile #	Raw			
	Square root			
	Log10			
	Inverse			
Duchenne Index	Raw			
	Square root			
	Log10			
	Inverse			
INTENSITY		No normal distributions		
AU6 Intensity	Raw			
	Square root			
	Log10			
	Inverse			
Duchenne AU12 Intensity	Raw			
	Square root			
	Log10			
	Inverse			

Non-Duchenne AU12 Intensity	Raw			
	Square root			
	Log10			
	Inverse			
Overall AU12 Intensity	Raw			
	Square root			
	Log10			
	Inverse			
ONSET				
DURATION				
AU12 First Peak Minus Onset	Raw			
	Square root	3	3	3
	Log10	3	3	3
	Inverse			

Note. First, normality was assessed using the Shapiro-Wilk test. If the distribution was normal in both the HC and PD group, homogeneity of variance was assessed with the Levene test. If homogeneity of variance was present, outliers were examined.

Table 8

Summary of Normality Testing for Global Measures of Facial Expressivity: Treatment by Time

Variable	Transformation	Treatment Group	Normality p-values 3 = Normal across treatment groups T1 (Baseline)	T2 (Post)	T3 (Follow-Up)	Homogeneity Of Variance?	Outliers Present? (If yes, outliers are specified: [#] = Subject #; T[1, 2, or 3]] = Time)
Lability	Raw		Sig.	Sig	Sig.	No(T2)	
		HC	.25	.96	.94		42@T1
		Untreated PD	<.01	.01	<.01		25@T1, 8@T3
		Artic	.01	.77	.01		26@T1, 39@T3
		LSVT	.02	.04	.02		none
	Square root			3[1]	3	3	
		HC	.78	.31	.75		50@T123
		Untreated PD	.03	.05	.07		25@T1 8@T3
		ARTIC	.41	.88	.78		none
		LSVT	.37	.31	.14		none
	Log10		3		3	3	
		HC	.31	.02	.38		50@T123
		Untreated PD	.66	.06	.45		25@T1 33@T1
		ARTIC	.52	.52	.72		none
		LSVT	.17	.18	.32		none

	Inverse					
		HC	<.01	<.01	.02	
		Untreated PD	.01	.01	.01	
		ARTIC	<.01	<.01	<.01	
		LSVT	<.01	<.01	.02	
Variability	Raw			3		No(T2)
		HC	.40	.74	.79	none
		Untreated PD	.49	.14	.02	8@T3
		ARTIC	.01	.52	.19	26@T1
		LSVT	.23	.18	<.01	5@T3
	Square root		3	3		3
		HC	.47	.76	.99	none
		Untreated PD	.27	.43	.35	8@T3
		ARTIC	.81	.15	.82	none
		LSVT	.06	.24	.04	65@T2, 5@T3
	Log10		3	3	3	3
		HC	.22	.50	.94	none
		Untreated PD	.09	.44	.51	none
		ARTIC	.97	.14	.59	none
		LSVT	.05	.14	.37	65@T2
	Inverse					
		HC	<.01	.02	.09	
		Untreated PD	<.01	.07	.09	
		ARTIC	.02	<.01	<.01	
		LSVT	<.01	<.01	.53	
Complexity	Raw					3
		HC	.01	.68	.43	42,14@T1
		Untreated PD	.03	.02	.42	none

	ARTIC	.49	.19	.01	58@T1,3
	LSVT	.52	.05	.58	27, 62, 65@T2
Square root					
	HC	.03	.77	.47	
	Untreated PD	.07	.07	.63	
	ARTIC	<.01	.02	<.01	
	LSVT	.01	.02	.48	
Log10			3		
	HC	.04	.75	.47	
	Untreated PD	.07	.10	<.01	
	ARTIC	.03	.08	<.01	
	LSVT	.06	.10	.32	
Inverse					
	HC	.24	.57	.49	
	Untreated PD	.08	.25	<.01	
	ARTIC	<.01	<.01	<.01	
	LSVT	<.01	<.01	.19	

Note. Untransformed scores were examined for normality, homogeneity of variance, and outliers. For transformed scores, first, normality was examined with the Shapiro-Wilk test. If a distribution did not significantly deviate from normal, homogeneity of variance was assessed with the Levene test. If variance was homogenous, distributions were examined for outliers.
[1]Normal across groups, with the exception of a *p*-value of .05 within the PD-Untreated group.

Table 9

Summary of Normality Testing for Measures of Smile Behavior: Treatment by Time

Variable	Transformation	Treatment Group	Normality p-values 3 = Normal across treatment groups T1 (Baseline)	T2 (Post)	T3 (Follow-Up)	Homogeneity Of Variance?	Outliers Present? (If yes, outliers are specified: [#] = Subject #; T[1, 2, or 3]] = Time)
FREQUENCY Duchenne #	Raw						
		HC	<.01	.02	.02		14, 42@T2
		Untreated PD	<.01	<.01	<.01		8, 53, 73@T1; 8, 53, 31@ T2; 8, 73@T3
		ARTIC	<.01	<.01	<.01		26@T1; 16, 55@T2; 39@T3
		LSVT	<.01	<.01	<.01		6, 63@T1; 5,24,71@T2; 24,34@T3
	Square root						
		HC	<.01	.03	.01		
		Untreated PD	<.01	<.01	<.01		

		ARTIC	<.01	<.01	<.01	
		LSVT	<.01	<.01	<.01	
	Log10					
		HC	<.01	.04	.01	
		Untreated PD	<.01	<.01	<.01	
		ARTIC	<.01	<.01	<.01	
		LSVT	<.01	<.01	<.01	
	Inverse					
		HC	<.01	.01	<.01	
		Untreated PD	<.01	<.01	<.01	
		ARTIC	<.01	<.01	<.01	
		LSVT	<.01	<.01	<.01	
Non-Duchenne #	Raw					
		HC	.08	.14	.30	none
		Untreated PD	<.01	.03	<.01	73@T2 8, 60@T3
		ARTIC	.15	.18	.04	none
		LSVT	<.01	.02	.25	5@T1 6, 71@T2
	Square root					
		HC	.03	.03	.28	
		Untreated PD	<.01	.02	.02	
		ARTIC	.06	.16	.10	
		LSVT	<.01	.03	.08	
	Log10					
		HC	.31	.06	.22	
		Untreated PD	<.01	.04	.02	

		ARTIC	.10	.25	.07	
		LSVT	<.01	.06	.15	
	Inverse					
		HC	.02	.01	<.01	
		Untreated PD	<.01	<.01	<.01	
		ARTIC	.01	.01	<.01	
		LSVT	<.01	<.01	<.01	
Total Smile #	Raw					
		HC	.42	.58	.83	none
		Untreated PD	<.01	.09	.01	8, 60@T3
		ARTIC	.19	.04	.01	55@T2 39@T3
		LSVT	<.01	<.01	.23	5, 63@T1 6, 71@T2
	Square root				3	
		HC	.05	.59	.88	
		Untreated PD	<.01	.02	.09	
		ARTIC	.18	.31	.19	
		LSVT	.01	.01	.13	
	Log10				3	
		HC	.05	.68	.66	
		Untreated PD	<.01	.04	.11	
		ARTIC	.20	.38	.16	
		LSVT	.01	.02	.12	
	Inverse					
		HC	<.01	.01	<.01	
		Untreated PD	<.01	<.01	.01	
		ARTIC	<.01	.01	<.01	

		LSVT	<.01	<.01	<.01	
Duchenne Index	Raw				
		HC	.01	.09	.06	10@T3
		Untreated PD	<.01	<.01	<.01	none
		ARTIC	<.01	<.01	.01	none
		LSVT	<.01	<.01	<.01	5, 24, 65@T2
	Square root				
		HC	<.01	.12	.23	
		Untreated PD	<.01	<.01	<.01	
		ARTIC	<.01	<.01	.01	
		LSVT	<.01	<.01	<.01	
	Log10				
		HC	<.01	.12	.31	
		Untreated PD	<.01	<.01	<.01	
		ARTIC	<.01	<.01	.01	
		LSVT	<.01	<.01	<.01	
	Inverse				
		HC	.08	.16	.71	
		Untreated PD	<.01	.006	<.01	
		ARTIC	.09	<.01	.01	
		LSVT	<.01	.27	<.01	
INTENSITY						
AU6 Intensity	Raw				
		HC	<.01	.11	.16	none
		Untreated PD	<.01	<.01	<.01	8, 31, 53@T2; 8, 53, 73@T3
		ARTIC	<.01	<.01	.01	16, 55@T2
		LSVT	<.01	<.01	<.01	6, 24, 63@T1

	Square root				
		HC	<.01	.01	.02
		Untreated PD	<.01	<.01	<.01
		ARTIC	<.01	<.01	<.01
		LSVT	<.01	<.01	<.01
	Log10				
		HC	<.01	02	.03
		Untreated PD	<.01	<.01	<.01
		ARTIC	<.01	<.01	<.01
		LSVT	<.01	<.01	<.01
	Inverse				
		HC	<.01	<.01	<.01
		Untreated PD	<.01	<.01	<.01
		ARTIC	<.01	<.01	<.01
		LSVT	<.01	<.01	<.01
Duchenne AU12 Intensity	Raw				
		HC	<.01	.02	.08	none
		Untreated PD	<.01	<.01	<.01	8, 48, 53, 73@T1; 31, 53@T2; 8, 48, 53@T3
		ARTIC	<.01	<.01	<.01	19@T1; 16, 55@T2
		LSVT	<.01	<.01	.01	6, 63@T1; 5, 24, 71@T2
	Square root				
		HC	<.01	<.01	.01	

		Untreated PD	<.01	<.01	<.01	
		ARTIC	<.01	<.01	<.01	
		LSVT	<.01	<.01	<.01	
	Log10					
		HC	<.01	<.01	.01	
		Untreated PD	<.01	<.01	<.01	
		ARTIC	<.01	<.01	<.01	
		LSVT	<.01	<.01	<.01	
	Inverse					
		HC	<.01	<.01	<.01	
		Untreated PD	<.01	<.01	<.01	
		ARTIC	<.01	<.01	<.01	
		LSVT	<.01	<.01	<.01	
Non-Duchenne AU12 Intensity	Raw					
		HC	.01	.01	.25	13, 17 @T1
		Untreated PD	.01	.03	.04	none
		ARTIC	.39	.27	.03	none
		LSVT	<.01	.04	.03	none
	Square root (negative skew)					
		HC	.01	.01	.49	
		Untreated PD	.01	.06	.05	
		ARTIC	.43	.40	.04	
		LSVT	<.01	.07	.06	
	Log10 (negative skew)					
		HC	.03	.02	.67	

		Untreated PD	.01	.10	.05	
		ARTIC	.28	.49	.05	
		LSVT	<.01	.09	.10	
	Inverse (negative skew)					
		HC	.12	.03	.73	
		Untreated PD	<.01	.16	.03	
		ARTIC	.02	.44	.08	
		LSVT	<.01	.07	.21	
Overall AU12 Intensity	Raw					
		HC	.17	.03	.11	13, 17@T1; 13, 14@T2; 17@T3
		Untreated PD	<.01	.09	.08	none
		ARTIC	.35	.09	.03	none
		LSVT	.01	.02	.03	65@T2
	Square root (negative skew)				3	
		HC	.38	.06	.28	
		Untreated PD	<.01	.22	.11	
		ARTIC	.44	.16	.06	
		LSVT	.01	.03	.07	
	Log10 (negative skew)				3	
		HC	.49	.03	.55	
		Untreated PD	.01	.38	.12	
		ARTIC	.343	.27	.07	
		LSVT	.008	.04	.14	

	Inverse (negative skew)						
		HC	.16	<.01	.90		
		Untreated PD	<.01	.46	.07		
		ARTIC	.03	.43	.02		
		LSVT	.01	.04	.38		
ONSET							
DURATION							
AU12 First Peak	Raw						
		HC	.40	.16	-		50@T1
		Untreated PD	.73	.11	-		none
		ARTIC	.03	.86	-		55@T1
		LSVT	.21	<.01	-		63@T2
	Square root		3				
		HC	.89	.13	-		none
		Untreated PD	.65	.37	-		none
		ARTIC	.14	.90	-		none
		LSVT	.85	.03	-		63@T2
	Log10		3	3		3	
		HC	.96	.10	-		none
		Untreated PD	.40	.59	-		none
		ARTIC	.39	.67	-		58@T2
		LSVT	.94	.85	-		none
	Inverse		3				multiple
		HC	.09	.06			
		Untreated PD	.06	.04			
		ARTIC	.56	.07			
		LSVT	.05	.02			

Note. Untransformed scores were examined for normality, homogeneity of variance, and outliers. For transformed scores, first, normality was examined with the Shapiro-Wilk test. If a distribution did not significantly deviate from normality across at any time, homogeneity of variance was assessed with the Levene test. If variance was homogenous, distributions were re-examined for outliers.

Table 10

Correlations Between Beck Depression Inventory (BDI) Score and Global Facial Expressivity Variables

			Correlation								
			Overall			Men			Women		
			rho	*p*	*n*	*rho*	*p*	*n*	*rho*	*p*	*n*
Baseline Correlations by Group											
Lability	HC	Baseline	.108	.753	11	.632	.368	4	-.165	.723	7
	PD	Baseline	.264	.080	45	.245	.170	33	.083	.798	12
Variability	HC	Baseline	.153	.653	11	.894	.106	4	-.074	.875	7
	PD	Baseline	.166	.277	45	.178	.321	33	-.080	.805	12
Complexity	HC	Baseline	-.181	.594	11	.800	.200	4	-.459	.300	7
	PD	Baseline	.044	.772	45	.078	.666	33	-.347	.269	12
Correlations by Time and Treatment Group											
Lability	HC	Baseline	.108	.753	11	.632	.368	4	-.165	.723	7
		Post	.035	.919	11	1.000	<.01*	4	-.352	.439	7
		Follow-Up	-.143	.693	10	.200	.800	4	-.239	.648	6
	PD-Untreated	Baseline	.128	.625	17	.059	.849	13	-.600	.400	4
		Post	.155	.552	17	.335	.264	13	-.600	.400	4
		Follow-Up	.044	.872	16	-.021	.948	12	-.800	.200	4

	ARTIC	Baseline	.276	.386	12	.426	.293	8	-.316	.684	4
		Post	-.100	.756	12	-.176	.677	8	-.632	.368	4
		Follow-Up	-.684	.014*	12	-.647	.083	8	-.500	.500	4
	LSVT	Baseline	.357	.175	16	.388	.212	12	<.001	1.000	4
		Post	.302	.256	16	.011	.974	12	.800	.200	4
		Follow-Up	.214	.455	15	-.096	.779	11	.738	.262	4
Variability	HC	Baseline	.153	.653	11	.894	.106	4	-.074	.875	7
		Post	.084	.807	11	1.000	<.01*	4	-.202	.664	7
		Follow-Up	-.038	.918	10	.400	.600	4	-.303	.559	6
	PD-Untreated	Baseline	-.053	.840	17	-.055	.859	13	-.600	.400	4
		Post	.279	.278	17	.378	.203	13	-.600	.400	4
		Follow-Up	-.147	.586	16	-.429	.164	12	-.400	.600	4
	ARTIC	Baseline	.277	.384	12	.427	.292	8	-.316	.684	4
		Post	.009	.978	12	.012	.977	8	-.500	.500	4
		Follow-Up	-.611	.035*	12	-.577	.134	8	-.632	.368	4
	LSVT	Baseline	.316	.233	16	.357	.255	12	-.200	.800	4
		Post	.365	.164	16	.290	.360	12	.800	.200	4
		Follow-Up	.159	.571	15	-.047	.892	11	.800	.200	4
Complexity	HC	Baseline	-.181	.594	11	.800	.200	4	-.459	.300	7
		Post	-.427	.190	11	-.400	.600	4	-.330	.469	7
		Follow-Up	-.262	.465	10	.000	1.000	4	-.478	.338	6
	PD-	Baseline	-.069	.792	17	-.033	.914	13	-.600	.400	4

Untreated										
	Post	.283	.271	17	.685	.010**	13	-.800	.200	4
	Follow-Up	-.269	.314	16	-.307	.332	12	-.600	.400	4
ARTIC	Baseline	.014	.965	12	.245	.558	8	-.632	.368	4
	Post	-.404	.192	12	-.252	.548	8	-.949	.051	4
	Follow-Up	-.204	.526	12	-.108	.799	8	-.949	.051	4
LSVT	Baseline	.257	.336	16	.322	.308	12	-.400	.600	4
	Post	.195	.470	16	.291	.358	12	.000	1.000	4
	Follow-Up	.119	.674	15	-.263	.435	11	.800	.200	4

Table 11

Correlations Between Beck Depression Inventory (BDI) Score and Smile Behavior Variables

			Correlation (Spearman)		
			rho	*p*	*N*
Baseline Correlations by Group					
FREQUENCY					
Duchenne #	HC	Baseline	.237	.482	11
	PD	Baseline	.142	.353	45
Non-Duchenne #	HC	Baseline	-.057	.868	11
	PD	Baseline	.007	.964	45
Total Smile #	HC	Baseline	.007	.984	11
	PD	Baseline	.061	.689	45
Index	HC	Baseline	.082	.810	11
	PD	Baseline	.134	.381	45
INTENSITY					
AU 6 Intensity	HC	Baseline	.070	.838	11
	PD	Baseline	.038	.804	45
Duchenne AU12 Intensity	HC	Baseline	.237	.482	11
	PD	Baseline	.089	.563	45
Non-Duchenne AU12 Intensity	HC	Baseline	.460	.155	11
	PD	Baseline	-.178	.243	45

Overall AU 12 Intensity	HC	Baseline	.404	.218	11
	PD	Baseline	-.027	.863	45
ONSET DURATION[1]					
First Peak AU12	HC	Baseline	.038	.922	9
	PD	Baseline	-.006	.978	9
Correlations by Time and Treatment Group					
FREQUENCY					
Duchenne #	HC	Baseline	.237	.482	11
		Post	.151	.657	11
		Follow-Up	-.251	.484	10
	PD-Untreated	Baseline	.485	.049*	17
		Post	.564	.018*	17
		Follow-Up	.222	.409	16
	ARTIC	Baseline	-.102	.753	12
		Post	-.049	.881	12
		Follow-Up	-.642	.024*	12
	LSVT	Baseline	<.001	1.000	16
		Post	.088	.747	16
		Follow-Up	-.121	.667	15
Non-Duchenne #	HC	Baseline	-.057	.868	11
		Post	-.579	.062	11

		Follow-Up	.203	.575	10
	PD-Untreated	Baseline	-.051	.845	17
		Post	-.266	.302	17
		Follow-Up	-.096	.722	16
	ARTIC	Baseline	.048	.881	12
		Post	-.237	.459	12
		Follow-Up	-.248	.437	12
	LSVT	Baseline	.239	.373	16
		Post	.257	.336	16
		Follow-Up	-.277	.318	15
Total Smile #	HC	Baseline	.007	.984	11
		Post	-.229	.499	11
		Follow-Up	.119	.744	10
	PD-Untreated	Baseline	.052	.843	17
		Post	.006	.983	17
		Follow-Up	.005	.986	16
	ARTIC	Baseline	.016	.960	12
		Post	-.138	.670	12
		Follow-Up	-.409	.186	12
	LSVT	Baseline	.361	.169	16
		Post	.377	.151	16
		Follow-Up	-.283	.306	15

Duchenne Index	HC	Baseline	.082	.810	11
		Post	.325	.330	11
		Follow-Up	-.569	.086	10
	PD-Untreated	Baseline	.424	.090	17
		Post	.518	.033*	17
		Follow-Up	.318	.231	16
	ARTIC	Baseline	.180	.576	12
		Post	.244	.445	12
		Follow-Up	-.107	.741	12
	LSVT	Baseline	-.258	.335	16
		Post	.032	.906	16
		Follow-Up	.007	.979	15
INTENSITY					
AU6 Intensity	HC	Baseline	.070	.838	11
		Post	-.112	.743	11
		Follow-Up	.074	.840	10
	PD-Untreated	Baseline	.139	.594	17
		Post	.542	.025*	17
		Follow-Up	.281	.292	16
	ARTIC	Baseline	.199	.536	12
		Post	<.001	1.000	12
		Follow-Up	-.499	.099	12

	LSVT	Baseline	-.188	.486	16
		Post	-.148	.584	16
		Follow-Up	-.009	.976	15
Duchenne AU12 Intensity	HC	Baseline	.237	.482	11
		Post	.110	.749	11
		Follow-Up	-.155	.669	10
	PD-Untreated	Baseline	.295	.250	17
		Post	.511	.036*	17
		Follow-Up	.208	.439	16
	ARTIC	Baseline	-.088	.787	12
		Post	-.049	.881	12
		Follow-Up	-.753	.005*	12
	LSVT	Baseline	.033	.903	16
		Post	.088	.747	16
		Follow-Up	.183	.514	15
Non-Duchenne AU12 Intensity	HC	Baseline	.460	.155	11
		Post	-.017	.961	11
		Follow-Up	.144	.692	10
	PD-Untreated	Baseline	-.207	.425	17
		Post	-.002	.994	17
		Follow-Up	-.221	.410	16

	ARTIC	Baseline	-.592	.043*	12
		Post	-.418	.176	12
		Follow-Up	-.061	.850	12
	LSVT	Baseline	.310	.243	16
		Post	.103	.704	16
		Follow-Up	-.171	.542	15
Overall AU12 Intensity	HC	Baseline	.404	.218	11
		Post	.660	.027*	11
		Follow-Up	.363	.303	10
	PD-Untreated	Baseline	-.049	.852	17
		Post	.221	.394	17
		Follow-Up	-.050	.855	16
	ARTIC	Baseline	-.550	.064	12
		Post	.362	.247	12
		Follow-Up	-.601	.039*	12
	LSVT	Baseline	.524	.037*	16
		Post	.276	.301	16
		Follow-Up	.158	.575	15
ONSET DURATION[1]	HC	Baseline	.038	.922	9
		Post	-.243	.529	9
		Follow-Up			
	PD-	Baseline	.305	.506	7

Untreated				
	Post	.157	.626	12
ARTIC	Baseline	-.030	.944	8
	Post	-.025	.958	7
LSVT	Baseline	-.172	.683	8
	Post	.039	.894	14
	Follow-Up			

Note. [1]Pearson's *r* calculation derived from log10-transformed Onset Duration data and natural-log-transformed BDI score.

Table 12

Baseline Gender Differences in Dependent Variables: PD versus Healthy Controls

	PD	Healthy Controls
GLOBAL EXPRESSIVITY		
Lability	.02	.26
	W>M	
Variability	.03	.32
	W>M	
Complexity	.02	.21
	W>M	
SMILE BEHAVIOR		
Duchenne #	.03	.08
	W>M	W>M
Non-Duchenne #	.37	.50
	W>M	
Total Smile #	.21	.26
Duchenne Index	.440	.230
AU6 Intensity	.11	.08
		W>M
Duchenne AU12 Intensity	.07	.08
	W>M	W>M
Non-Duchenne AU12 Intensity	.31	.04
		W>M
Overall AU12 Intensity	.07	.04
	W>M	W>M
Onset Duration[1]	.16	.59
	M>W	

Note.All results obtained with one-tailed Mann-Whitney tests, with the exception of Duchenne Index, for which a two-tailed Mann-Whitney test was used, and Onset Duration, for which a two-tailed *t*-test was used.

[1] t-test on log10-transformed data).

Table 13

Change-Scores of Global Measures of Facial Expressivity: Means, Medians, and Standard Deviations of Change-Scores for Lability, Variability, and Complexity

		T2-T1			T3-T2			T3-T1		
		M	*SD*	*Mdn*	*M*	*SD*	*Mdn*	*M*	*SD*	*Mdn*
Lability	HC	-1.18	7.77	0.00	4.20	7.6	4.50	4.70	9.52	5.50
	PD-Untreated	-1.71	6.82	-1.00	4.56	8.67	3.00	2.69	9.67	1.00
	Artic	-1.25	6.68	0.50	7.92	13.32	4.00	6.67	13.68	2.50
	LSVT	5.00	8.24	2.50	2.93	8.42	3.00	6.93	6.44	6.00
Variability	HC	-0.73	2.64	0.00	3.80	3.79	3.00	3.30	4.92	1.50
	PD-Untreated	-0.94	2.33	-1.00	3.06	3.47	2.50	2.00	2.88	1.50
	ARTIC	-0.58	3.15	0.50	3.42	5.32	3.00	2.83	5.04	1.00
	LSVT	1.81	3.04	1.50	2.53	3.4	2.00	3.93	4.57	3.00
Complexity	HC	0.47	0.73	0.35	-0.86	1.08	-0.80	-0.33	1.16	0.14
	PD-Untreated	-0.09	0.60	-0.13	-0.29	0.75	-0.18	-0.37	0.63	-0.28
	ARTIC	0.11	0.75	0.13	-0.03	0.77	0.09	0.08	0.54	0.09
	LSVT	0.40	1.31	0.06	-0.49	1.08	-0.14	-0.07	0.94	0.04

Table 14

Smile Behavior: Means, medians, and standard deviations of raw scores

		T1			T2			T3		
		M	*SD*	*Mdn*	*M*	*SD*	*Mdn*	*M*	*SD*	*Mdn*
FREQUENCY[1]										
Duchenne #	HC	0.73	1.10	0.00	1.09	1.14	1.00	1.10	0.88	1.00
	PD-Untreated	0.29	0.69	0.00	0.24	0.56	0.00	0.44	0.89	0.00
	ARTIC	0.33	0.65	0.00	0.50	1.24	0.00	1.42	2.23	0.00
	LSVT	0.13	0.34	0.00	0.25	0.58	0.00	0.87	1.06	1.00
Non-Duchenne #	HC	2.36	2.11	2.00	1.36	1.12	1.00	3.60	3.06	3.00
	PD-Untreated	1.00	1.12	1.00	1.41	1.23	1.00	1.31	1.40	1.00
	ARTIC	1.75	1.48	1.50	1.75	1.60	1.50	3.08	3.32	3.00
	LSVT	0.81	1.11	0.50	1.50	1.26	1.00	1.87	1.51	2.00
Total Smile #	HC	3.09	2.17	3.00	2.45	1.75	2.00	4.70	3.50	4.50
	PD-Untreated	1.29	1.57	1.00	1.65	1.22	1.00	1.75	1.73	1.00
	ARTIC	2.08	1.93	1.50	2.25	2.34	2.00	4.50	4.93	4.00
	LSVT	0.94	1.18	1.00	1.75	1.48	1.00	2.73	2.28	3.00
Duchenne Index	HC	-0.43	0.57	-0.20	-0.17	0.71	0.00	-0.41	0.59	-0.55
	PD-Untreated	-0.36	0.49	0.00	-0.57	0.64	-1.00	-0.54	0.61	-1.00
	ARTIC	-0.60	0.45	-0.75	-0.56	-0.56	-1.00	-0.27	0.62	0.00
	LSVT	-0.40	0.61	-0.17	-0.65	0.70	-1.00	-0.39	0.41	-0.33

INTENSITY[2]										
AU6 Intensity	HC	1.03	1.58	0.00	1.60	1.48	2.00	1.80	1.48	2.00
	PD-Untreated	1.15	1.87	0.00	0.37	0.83	0.00	0.56	1.09	0.00
	ARTIC	0.79	1.20	0.00	0.42	0.97	0.00	1.22	1.35	0.72
	LSVT	0.63	1.26	0.00	0.84	1.36	0.00	1.36	1.41	1.33
Duchenne AU12 Intensity	HC	1.30	1.90	0.00	2.33	1.97	3.00	2.05	1.95	2.25
	PD-Untreated	0.94	1.85	0.00	0.38	1.08	0.00	0.75	1.39	0.00
	ARTIC	0.79	1.47	0.00	0.46	1.08	0.00	1.04	1.55	0.00
	LSVT	0.31	0.87	0.00	0.59	1.28	0.00	1.44	1.54	1.00
Non-Duchenne AU12 Intensity	HC	2.05	1.09	2.33	1.83	1.27	2.00	1.77	1.07	1.90
	PD-Untreated	1.30	1.38	1.33	1.61	1.08	2.00	1.53	1.13	1.90
	ARTIC	1.73	1.31	2.00	1.44	1.02	1.50	1.24	1.01	1.50
	LSVT	1.02	1.13	0.50	1.76	1.08	2.00	1.47	0.89	2.00
Overall AU12 Intensity	HC	2.19	1.27	2.33	2.74	1.11	3.00	2.14	0.92	2.48
	PD-Untreated	1.49	1.56	2.00	1.80	1.07	2.00	1.71	1.10	2.00
	ARTIC	1.77	1.32	2.00	1.56	1.05	1.75	1.73	1.15	2.00
	LSVT	1.21	1.20	1.33	2.17	0.90	2.29	1.63	0.97	2.00
ONSET DURATION[3] AU12 Onset Duration	HC	0.49	0.29	0.38	0.55	0.20	0.66	-	-	-

(Time to First Peak)										
	PD-Untreated	0.67	0.35	0.65	0.45	0.32	0.29	-	-	-
	ARTIC	0.43	0.24	0.39	0.37	0.15	0.35	-	-	-
	LSVT	0.58	0.48	0.48	0.46	0.51	0.27	-	-	-

Note.
[1]Frequency are measured as counts of the variable described.
[2]Average intensity based on the FACS Coding System.
[3]Onset Duration is measured in seconds.

Table 15

Baseline Differences in Global Measures of Expressivity. Nonparametric Evaluation of Untransformed Scores

	Group Means					
	HC	PD	*U*	*z*	*p*	Direction
ALL PARTICIPANTS *N*: HC = 11; PD = 45						
Lability	15.27	7.27	122.50	-2.59	.01*	HC>PD
Variability	7.00	3.80	145.50	-2.12	.02*	HC>PD
Complexity	1.83	1.51	215.50	-0.66	.25	
WOMEN *N*: HC = 7; PD = 12						
Lability	17.71	10.50	27.50	-1.23	.11	
Variability	7.86	5.17	32.50	-0.81	.22	
Complexity	2.10	1.93	43.00	0.09	.50	
MEN *N*: HC = 4; PD = 33						
Lability	11.00	6.09	20.00	-2.26	.01*	HC>PD
Variability	5.50	0.30	28.00	-1.87	.03*	HC>PD
Complexity	1.37	1.36	63.00	-0.15	.45	

Note. One-tailed Mann-Whitney tests comparing baseline global expressivity between Healthy Control (HC) and Parkinson's disease (PD) groups.

Table 16

Results of one-way Kruskal-Wallis tests examining change-scores (separate, for Lability, Variability, and Complexity) across Treatment Groups

	T2-T1 change-score		T3-T2 change-score		T3-T1 change-score	
	$\chi^2(3, N=56)$	*p*	χ^2 (3, *N*=53)	*p*	χ^2 (3, *N*=53)	*p*
ALL PARTICIPANTS						
Lability	7.28	.06* LSVT>UPD (.01) LSVT>ARTIC (.06) LSVT>HC(.09)	.53	.91	3.39	.34
Variability	8.01	.05** LSVT>UPD (.01)	1.07	.78	2.12	.55
Complexity	4.21	.24	3.89	.27	2.98	.39
WOMEN	X2(3, N=19)		X2(3, N=19)		X2(3, N=18)	
Lability	1.47	.69	.64	.89	.23	.97
Variability	0.66	.88	1.19	.76	.66	.88
Complexity	4.56	.21	2.09	.55	.52	.91
MEN	X2(3, N=37)		X2(3, N=37)		X2(3, N=35)	
Lability	6.02	.11	.48	.92	5.36	.15
Variability	6.37	.10 LSVT>UPD (.01)	1.04	.79	4.42	.22

	7.12	.07	6.11	.11	5.15	.16
Complexity		LSVT>UPD(.06); HC>UPD (.04); LSVT>ARTIC (.10); HC>ARTIC (.06)				

Note. For pairwise comparisons, unadjusted p-values are reported. UPD = Untreated PD group.

Table 17

Baseline Smile Behavior: All Healthy Controls Compared to All Individuals with PD

	Group Mean							
	HC	PD		*U*	*z*	*p*		Test
Frequency	0.73	0.24	Duchenne Number	194.00	-1.54	.06	HC>PD	one-tailed Mann-Whitney
	2.36	1.13	Non-Duchenne Number	158.50	-1.91	.06	HC>PD	two-tailed Mann-Whitney
	3.09	1.38	Total Smile Number	133.00	-2.44	.02*	HC>PD	two-tailed Mann-Whitney
	-0.43	-0.44	Duchenne Index	244.00	-0.08	.47		one-tailed Mann-Whitney
Intensity	1.03	0.87	AU6	229.00	-0.47	.32		one-tailed Mann-Whitney
	1.30	0.68	Duchenne AU12	205.00	-1.19	.12		one-tailed Mann-Whitney
	2.05	1.31	Non-Duchenne AU12	160.50	-1.85	.03*	HC>PD	one-tailed Mann-Whitney
	2.19	1.47	AU12-All	175.00	-1.54	.06	HC>PD	one-tailed Mann-Whitney
Onset Duration	0.49	0.56		*F*(1,28)	Partial η2	*p*		ANCOVA[1]
				2.33	.08	.14		
				t(30)		*p*		*t*-test[2]
				-0.34		.74		

Note.
HC: Healthy Controls. PD: Parkinson's disease.
[1]*DV=log10-transformed AU12 Onset Duration; Covariate = AU12 Intensity.. Single outlier excluded from analysis.*
[2]log10-transformed (equal variances assumed)
* Significant at alpha of .05. Relative size of group means is indicated for statistically significant and trend-level results.

Table 18

Gender Effects: Baseline Smile Behavior of Women Healthy Controls Compared to Women with PD

	Group Mean						
	HC	*PD*		*U*	*z*	*p*	Test
WOMEN			*N*: HC=7; PD = 12				
Frequency	1.14	0.58	Duchenne Number	30.50	-1.09	.17	one-tailed Mann-Whitney
	2.29	1.33	Non-Duchenne Number	29.00	-1.13	.30	two-tailed Mann-Whitney
	3.43	1.92	Total Smile Number	26.00	-1.38	.20	two-tailed Mann-Whitney
	-0.24	-0.29	Duchenne Index	38.00	-0.36	.39	one-tailed Mann-Whitney
Intensity	1.61	1.33	AU6	37.50	-0.41	.36	one-tailed Mann-Whitney
	2.05	1.29	Duchenne AU12	34.00	-0.75	.27	one-tailed Mann-Whitney
	2.33	1.51	Non-Duchenne AU12	28.00	-1.21	.13	one-tailed Mann-Whitney
	2.56	1.97	AU12-All	35.00	-0.60	.35	one-tailed Mann-Whitney
Onset Duration	0.55	0.45	*N*: HC = 6; PD = 9	*F*(1,9)	partial η2	*p*	ANCOVA[1]
				0.17	.02	.69	
				t(10)			*t*-test[2]
				.71		.50	
MEN			*N*: HC = 4; PD = 33	*U*	*z*	*p*	Test
Frequency	0.00	0.12	Duchenne Number	58.00	-0.73	.36	one-tailed Mann-Whitney
	2.50	1.06	Non-Duchenne Number	43.50	-1.15	.28	two-tailed Mann-Whitney
	2.50	1.18	Total SmileNumber	45.500	-1.05	.33	two-tailed Mann-Whitney
	-0.75	-0.49	Duchenne Index	48.00	-.98	.20	one-tailed Mann-Whitney

Intensity	0.00	0.70	AU6	50.00	-1.09	.23	one-tailed Mann-Whitney
	0.00	0.45	Duchenne AU12	56.00	-.82	.33	one-tailed Mann-Whitney
	1.54	1.24	Non-Duchenne AU12	60.00	-.31	.40	one-tailed Mann-Whitney
	1.54	1.29	AU12-All	61.50	-.23	.42	one-tailed Mann-Whitney
Onset Duration	0.36	0.59	*N*: HC = 3; PD = 24	*F*(1,17)	partial η2	p	ANCOVA[1]
				0.92	.05	.35	
				t(18)			
				-1.10		.29	*t*-test[2]

Note.

HC: Healthy Controls. PD: Parkinson's disease.
[1]*ANCOVAs conducted with log10-transformed AU 12 Onset Duration as DV and AU12 Intensity as covariate. No outliers.*
[2]Log10-transformed data. Equal variances assumed. As with the overall data set, when analyzed by gender, data was normally distributed and exhibited no outliers. Exact significance levels are reported for gender-specific analyses.

Table 19

Mean Baseline Scores for Smile Behavior Measures: All Participants, Women, and Men

Variable	All Participants		Women		Men	
	HC (n= 11)	PD (n = 45)	HC (n = 7)	PD (n = 12)	HC (n = 4)	PD (n = 33)
FREQUENCY						
Duchenne #[1]	0.73	0.24	1.14	0.58	0.00	0.12
Non-Duchenne #	2.36	1.13	2.29	1.33	2.25	1.06
Total Smile #	3.09	1.38	3.43	1.92	2.50	1.18
Duchenne Index	-0.43	-0.44	-0.24	-0.29	-0.75	-0.49
INTENSITY[1]						
AU6 Intensity	1.03	0.87	1.61	1.33	0.00	0.70
Duchenne AU12 Intensity	1.30	0.68	2.05	1.29	0.00	-0.45
Non-Duchenne AU12 Intensity	2.05	1.31	2.33	1.51	1.54	1.24
Overall AU12 Intensity	2.19	1.47	2.56	1.97	1.54	1.29
ONSET DURATION[2]	n = 9	n = 23	n = 6	n = 9	n = 3	n = 24
Time to First Peak AU12 (sec)	0.49	0.56	0.55	0.45	0.36	0.59

Note.
[1] One-tailed Mann-Whitney Test used to detect differences between baseline HC and PD scores.HC versus PD baseline means assessed via an ANCOVA with log10-transformed onset duration as the DV and log10-transformed AU12 Intensity as the covariate. *p*-value of main effect of group is provided.

Table 20

Impact of Treatment upon Smiling Behavior:Analysis of Change-Scores. (All Participants)

	T2 - T1		T3 - T2		T3 - T1	
	χ^2 (3, N=56)	*p*	χ^2 (3, N=53)	*p*	χ^2 (3, N=56)	*p*
Frequency[1]						
Duchenne Number	3.22	.36	2.03	.57	3.66	.30
Non-Duchenne Number	7.11	.07*	2.49	.48	1.15	.76
Total Smile Number	4.66	.20	2.55	.47	2.25	.52
Duchenne Index	4.47	.22	2.00	.57	3.77	.29
Intensity[1]						
AU6	3.56	.31	1.49	.69	7.11	.07*
Duchenne AU12	4.36	.23	1.86	.60	7.05	.07*
Non-Duchenne AU12	4.33	.23	0.33	.95	3.87	.28
AU12-All	5.69	.13	4.72	.19	2.21	.53
Onset Duration	ANOVA[2]					
	Time	$F(1, 25) = 0.08$, $p = .78$, partial $\eta 2 = <.01$				
	Treatment	$F(3, 25) = 0.47$, $p = .71$, partial $\eta 2 = .05$				
	Interaction	$F(3, 25) = 0.46$, $p = .72$, partial $\eta 2 = .05$				
	Kruskal-Wallis[1]					
	χ^2 (3, *N*=29)	*p*				

	3.10	0.38

Note. Planned contrasts were performed if omnibus test statistic was significant. Remaining pairwise comparisons reported on exploratory basis. See text.
[1]Results of Kruskal-Wallis.
[2]Results of ANOVA (2[Time]] x 4[Treatment Group]) using log10-transformed data.

Table 21

Smiling Behavior: Wilcoxon Signed-Rank Tests Comparing Time 1 Scores to Time 2 Scores

			z	*p*	direction
FREQUENCY	Duchenne Number	HC	1.03	.31	
		PD-Untreated	-0.58	.56	
		Artic	0.55	.58	
		LSVT	0.71	.48	
	Non-Duchenne Number	HC	-1.56	.12	
		PD-Untreated	1.15	.25	
		Artic	0.11	.92	
		LSVT	1.78	.07	T2>T1
	Total Smile Number	HC	-0.92	.36	
		PD-Untreated	0.98	.33	
		Artic	0.34	.73	
		LSVT	1.67	.10	T2>T1
	Duchenne Index	HC	0.92	.36	
		PD-Untreated	-1.54	.13	
		Artic	0.14	.89	
		LSVT	-1.04	.30	
INTENSITY	AU6	HC	0.95	.34	

	PD-Untreated	-1.79	.07	T2<T1
	Artic	-0.81	.42	
	LSVT	0.63	.53	
Duchenne AU12	HC	1.19	.23	
	PD-Untreated	-1.21	.23	
	Artic	-0.82	.41	
	LSVT	0.96	.34	
Non-Duchenne AU12	HC	-0.24	.81	
	PD-Untreated	1.20	.23	
	Artic	-1.05	.29	
	LSVT	1.71	.09	T2>T1
AU12-All	HC	1.68	.09	T2>T1
	PD-Untreated	0.87	.39	
	Artic	-0.49	.62	
	LSVT	2.51	.01	T2>T1

Figure 1

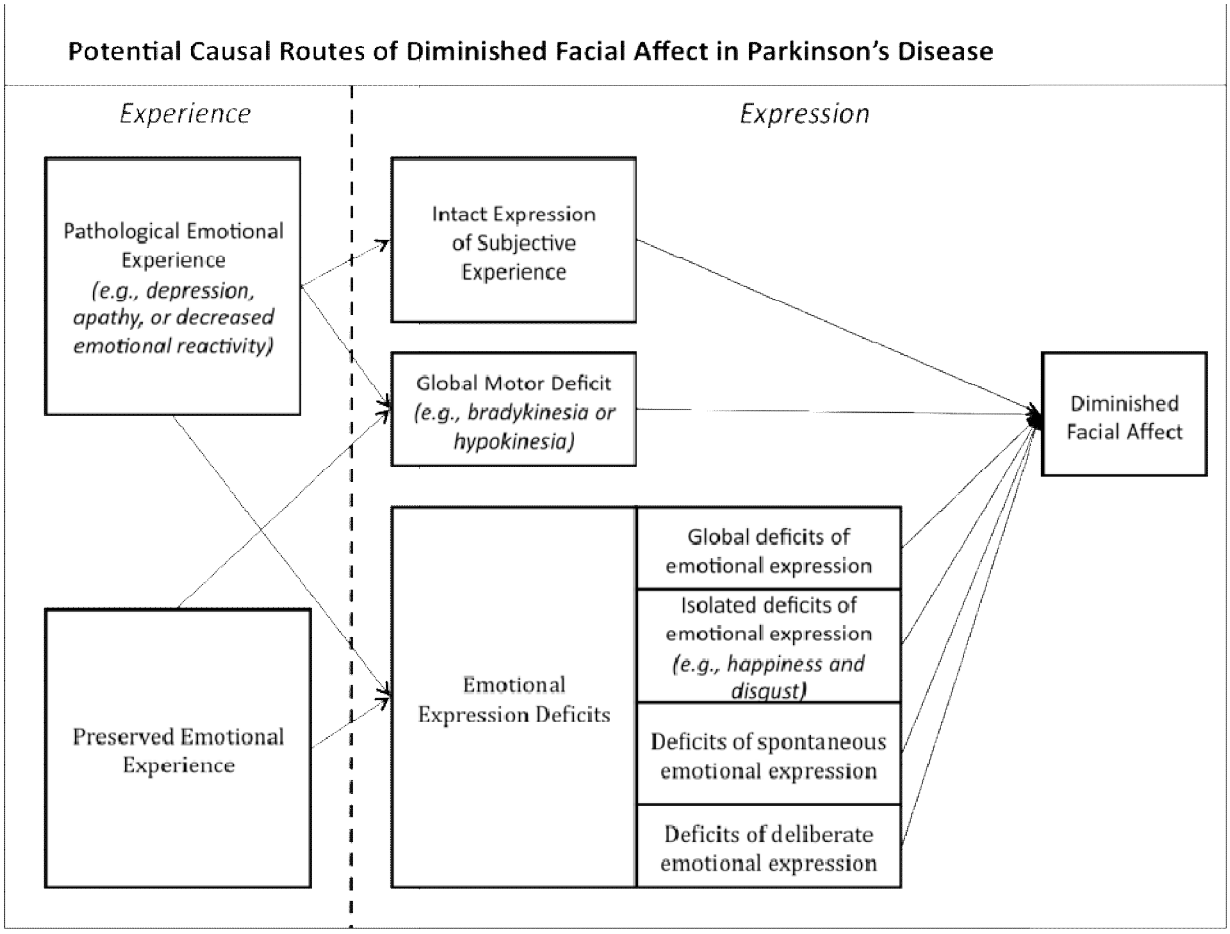

Figure 1. Potential causal routes resulting in diminished facial affect in Parkinson's disease. Several non-mutually exclusive possibilities are depicted. The direction of causality flows from left to right (from McCabe, D., Borod, J., Meltzer, E., Spielman, J., & Ramig, L., 2010).

References

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.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression.*Journal of Personality and Social Psychology, 70,* 614-636.

Baumgartner, C., Sapir, S., & Ramig, L. (2001). Voice quality changes following phonatory-respiratory effort Treatment (LSVT) versus respiratory effort treatment for individuals with Parkinson disease. *Journal of Voice, 15*, 105-114

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

Bleuler E. (1983). *Lehrbuch der Psychiatrie*, 15th ed. Berlin: Springer.

Borod, J.C. (1993a). Cerebral mechanisms underlying facial, prosodic, and lexical emotional expression: A review of neuropsychological studies and methodological issues. *Neuropsychology, 7,* 445-463.

Borod, J. C. (1993b). Emotion and the brain -- Anatomy and theory: An introduction to the Special Section. *Neuropsychology, 7,* 427-432.

Borod. J. C., & Brickman, A. (2001). Emotional processing deficits in Parkinson's disease: A review of the literature [Abstract]. *Journal of the International Neuropsychological Society, 7,* 243-244.

Borod, J. C., Alpert, M., Brozgold, A., Martin, C., Welkowitz, J., Diller, L., Peselow, E., Angrist, & Lieberman, A. (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., Halfacre, M., Rogers, K., Spielman, J., Krch, D., Brickman, A., Gruber, C., McCabe, D., & Ramig, L. (2007, February). *Aspects of facial emotion expression in Parkinson's disease and healthy control participants.* Paper presented at the International Neuropsychological Society, Portland, OR.

Borod, J.C.,Haywood, C.S. and Koff, E.(1997). Neuropsychological aspects of facial asymmetry during emotional expression: A review of the normal adult literature. *Neuropsychology Review, 7,* 41-60.

Borod, J. C., &Koff, E. (1984). Asymmetries in affective facial expression: Anatomy and behavior. In Fox, N., and Davidson, R. (Eds.), *The psychobiology of affective development*, (pp. 293-323). Lawrence Erlbaum Associates, Hillsdale: NJ.

Borod, J. C., & Madigan, N. (2000). Neuropsychology of emotion and emotional disorders: An overview and research directions. In J. Borod (Ed.), *The neuropsychology of emotion*. New York, NY: Oxford University Press.

Borod, J. C., Lorch, M. P., Koff, E., & Nicholas, M. (1987). Effect of emotional context onbucco-facial apraxia. *The Journal of Clinical and Experimental Neuropsychology, 9*, 155-161.

Borod, J. C., Rogers, K., Spielman, J., Halfacre, M., McCabe, D., Flanagan, T., & Ramig, L. (2008, June).*Emotional experience and expression in Parkinson's disease.* Paper presented at the Movement Disorders Society, Chicago, IL.

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

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

Bowers, D., Miller, K., Bosch, W., Gokcay, D., Pedraza, O., Springer, U., et al. (2006). 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., et al. (2006). Startling facts about emotion in Parkinson's disease: Blunted reactivity to aversive stimuli. *Brain, 129*, 3356-3365.

Bowers, D., Sapienza, C., Okun, M., & Fernandez, H. (Ongoing clinical trial).*Masked Faces in Parkinson Disease: Mechanism and Treatment.* Retrieved August 3, 2009, from ClinicalTrials.gov Web site: http://clinicaltrials.gov/ct2/show/NCT00350402

Breitenstein, C., Van Lancker, D., Daum, I., & Waters, C.H. (2001). Impaired perception of vocal emotions in Parkinson's disease: Influence of speech time processing and executive functioning. *Brain and Cognition*, *45*, 277-314.

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.

Brozgold, A., Borod, J., Rosen, J., & Alpert, M. (1999, August). *Facial emotional expression in patients with right-brain-damage and Parkinson's disease.* Paper presented at the American Psychological Association, Boston, MA.

Buck, R., & Duffy, R. J. (1980). Nonverbal communication of affect in brain-damaged

patients.*Cortex, 16*, 351-362.

CensusScope. (2009). *Demographic maps: An aging population*. Retrieved August 5, 2009, from CensusScope Web site: http://www.censusscope.org/us/map_65plus.html

Chung, S. J., Jeon, S. R., Kim, S. R., Sung, Y. H., & Lee, M. C. (2006). Bilateral effects of unilateral subthalamic nucleus deep brain stimulation in advanced Parkinson's disease. *European Neurology, 56*, 127-132.

Crucian, G.P., Huang, L., Barrett, A.M., Schwartz, R.L., Cibula, J.E., Anderson, J.M., Triggs, W.J., Bowers, D., Friedman, W.A., Greer, M., &Heilman, K.M. (2001). Emotional conversations in Parkinson's Disease. *Neurology*, *56*, 159-165.

Cummings, J. L. (1992). Depression in Parkinson's disease: A review. *The American Journal of Psychiatry, 149*, 443-454.

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

Cummings, J. L., Arciniegas, D. B., Brooks, B. R., Herndon, R. M., Lauterbach, E. C., Pioro, E. P., et al. (2006). Defining and diagnosing involuntary emotional expression disorder.*CNS Spectrums, 11*, 1-7.

de Lau, L. M., &Breteler, M. M. (2006). Epidemiology of Parkinson's disease.*Lancet Neurology, 5*, 525-535

Dickson, K. L., Walker, H., &Fogel, A. (1997). The relationship between smile type and play type during parent-infant play. *Developmental Psychology, 33*, 925-933.

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.

Drevets, W. C. (1998). Functional neuroimaging studies of depression: The anatomy of melancholia. *Annual Review of Medicine, 49*, 341-361.

Duchenne de Boulogne, C. B. (1990) *The Mechanism of Human Facial Expression* (R. Andrew Cuthbertson, Ed., Trans). Cambridge: Cambridge University Press. (Original work published 1862).

Dujardin, K., Defebvre, L., Duhamel, A., Lecouffe, P., Rogelet, P., Steinling, M., &Destée, A. (2004). Cognitive and SPECT characteristics predict progression of Parkinson's disease in newly diagnosed patients. *Journal of Neurology*, *251*, 1383-1392.

Dumer, A. I., Borod, J. C., Oster H., Spielman, J. L., Ramig, L. O., & Rabin, L. A. (2010,

November). *The effect of the Lee Silverman Voice Treatment (LSVT) on masked facies in Parkinson's disease: Preliminary data*. Paper presented at the Movement Disorders Conference, Moscow, Russia.

Dumer, A.I. (2011). *The effects of Lee Silverman Voice Treatment on the facial movement of Parkinson's disease patients and the way they are perceived.* (Unpublished doctoral dissertation). The Graduate Center of the City University of New York., New York

Eccles, J. (1980). The emotionalbrain. *Bulletin et Memoirs de L'AcademieRoyle de Medicine de Belgique, 135*, 697-711.

Ekman, P. (1989). The argument and evidence about universals in facial expressions of emotion. In H. Wagner & A. Manstead (Eds.), *Handbookof psychophysiology: The biological psychology of emotions andsocial processes* (pp. 143-164). London: John Wiley Ltd.

Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The Duchenne smile: Emotional expression and brain physiology II. *Journal of Personality and Social Psychology, 58,* 342-353.

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. (1975). *Unmasking the face: A guide to recognizing emotions from facial clues*. Oxford, England: Prentice-Hall.

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

Ekman, Friesen, Hager (2002). *Facial action coding system: The manual on CD ROM*. Salt Lake City, Utah: Research Nexus division of Network Information Research Corporation.

Ekman, P., Friesen, W. V., & O'Sullivan, M. (1988). Smiles when lying.*Journal of Personality and Social Psychology, 54,* 414-420.

Elefant, C., Lotan, M., Baker, F.A. &Skeie, G.O. (2012). Effects of music therapy on facial expression of individuals with Parkinson's disease: A pilot study. *Musicae Scientiae*, *16*, 392-400

Fagerstrom, K. O., Pomerleau, O., Giordani, B., &Stelson, F. (1994). Nicotine may relieve symptoms of Parkinson's disease. *Psychopharmacology, 116*, 117-119.

Fahn, S., Elton, R. L. (1987). UPDRS development committee. Unified Parkinson's Disease Rating Scale. In: S. Fahn, C .D. Marsden, D. B. Calne, & M. Goldstein (Eds.), *Recent Developments in Parkinson's Disease* (153-163). Florham Park, NJ: Macmillan.

Fahn, S. (2003).Description of Parkinson's disease as a clinical syndrome.*Annals of the New*

York Academy of Sciences, *991,* 1-14.

Fallon, J. H., Koziell, D. A., & Moore, R. Y. (2004). Catecholamine innervation of the basal forebrain II.Amygdala, suprarhinal cortex and entorhinal cortex.*The Journal of Comparative Neurology, 180*, 509-531.

Fernandez-Dols, J. M., & Ruiz-Belda, M. A. (1997). Spontaneous facial behavior during intense emotional episodes: Artistic truth and optical truth. In J. A. Russel& J. M. Ferna´ndez-Dols (Eds.), *The psychology of facial expression* (pp. 255-274). New York: Cambridge University Press.

Fox, C.M., Morrison, C.E., Ramig, L.O., & Sapir, S. (2002). Current perspectives on the Lee Silverman Voice Treatment (LSVT) for individuals with idiopathic Parkinson disease.*American Journal of Speech-Language Pathology, 11,* 111-123.

Fox, N. A., & Davidson, R. J. (1988). Patterns of brain electric activity during facial signs of emotion in 10-month-old infants.*Developmental Psychology, 24,* 230-236.

Fox, C.M., Ramig, L.O., Ciucci, M.R., Sapir, S., McFarland, D.H., & Farley, B.G. (2006). Neural plasticity—Principled approach to treating individuals with Parkinson disease and other neurological disorders.*Seminars in Speech and Language, 27,* 283-299.

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.

Fridlund, A. J. (1991). Sociality of solitary smiling: Potentiation by an implicit audience. *Journal of Personalityand Social Psychology, 60,* 229-240.

Fridlund, A. J. (1994). *Human facial expression: An evolutionary view.* San Diego, CA: Academic Press.

Frisina, P. G., Borod, J. C., Foldi, N. S., &Tenenbaum, H. R. (2008). Depression in Parkinson's disease: Health risks, etiology, and treatment options. *Neuropsychiatric Disease and Treatment, 4*, 81-89.

Frisina, P. G., Tenenbaum, H. R., Borod, J. C., & Foldi, N. (2008). The effects of antidepressants in Parkinson's disease: A meta-analysis. *International Journal of Neuroscience, 118,* 667-682.

Fudge, J. L. & Emiliano, A. B. (2003). The extended amygdala and the dopamine system: Another piece of the dopamine puzzle. *The Journal of Neuropsychiatry and Clinical Neurosciences, 15*, 306-316.

Galasko, D., Katzman, R., Salmon, D. P., Hansen, L. (1996).Clinical and neuropathological findings in Lewy Body dementia.*Brain and Cognition, 31*, 166-175.

Glozman, ZhM., Levin, OS, &Lycheva, Nlu (2003). Disorders of emotional memory and ability to identify emotional states in patients with Parkinson's disease.*FiziolCheloveka*, *6*, 55-60.

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.

Gray, H. M., & Tickle-Degnen, L. T. (2010). A meta-analysis of performance on emotion recognition tasks in Parkinson's disease. *Neuropsychology, 24,* 176-191.

Grunwald, I., Borod, J., Obler, L., Erhan, H., Pick, L., Welkowitz, J., Madigan, N., Sliwinski, M., &Whalen, J. (1999). The effects of age and gender on the perception of lexical emotion. *Applied Neuropsychology,6*, 226-238.

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

Halfacre, M., Dumer, A., Alterescu, K., McCabe, D., Ramig, L.O., Spielman, J.L., & Borod, J.C. (2009). Facial emotional expression in Parkinson's disease: Neural mechanisms. In A. Freitas-Magalhães (Ed.), *Emotional expression: The brain and the face* (pp. 95-126). Porto, Portugal: University Fernando Pessoa Press.

Hallgren, K.A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods Psychology.8,* 23-34.

Harding, A. J., Stimson, E., Henderson, J. M., &Halliday, G. M. (2002). Clinical correlates of selective pathology in the amygdala of patients with Parkinson's disease. *Brain, 125*, 2431-2445.

Heilman, K. M., Blonder, L. X., Bowers, D., &Crucian, G. P. (2000). Neurological disorders and emotional dysfunction. In J. C. Borod (Ed.), *The neuropsychology of emotion* (pp. 367-412). New York: Oxford University Press.

Heilman, K. M., Blonder, L. X., Bowers, D., &Valenstein, E. (2003). Emotional disorders associated with neurological diseases. In K. M. Heilman& E. Valenstein (Eds.), *Clinical neuropsychology* (pp. 447-478). New York: Oxford University Press.

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

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* (81-91). New York: DernosVermande.

Hohl, U., Tiraboschi, P., Hansen, L. A., Thal, L. J., & Corey-Bloom, J. (2000). Diagnostic accuracy of dementia with lewy bodies. *Archives of Neurology, 57*, 347-351.

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

Izard, C.E. (1983). *The maximally discriminative facial movement coding system (Max) (Revised)*. Newark, DE: Instructional Resources Center, University of Delaware.

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.

Jankovic, J. (2003). Pathophysiology and clinical assessment of parkinsonian symptoms and signs. In R. Pahwa, K. E. Lyons, & W. C. Koller (Eds.), *Handbook of Parkinson's disease*(3rd ed., pp. 71-107). New York: Marcel Dekker.

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.

Juckel, G., Mergl, R., ParBl, A., Mavrogiorgou,, P., Witthaus, H., Moller, H. J., Hegerl, U. (2008). Kinematic analysis of facial behavior in patients with schizophrenia under emotional stimulation by films with "Mr. Bean." *European Archives of Psychiatry and Clinical Neuroscience, 258*, 186-191.

Kan, Y., Kawamura, M., Hasegawa, Y., Mochizuki, S., & Nakamura, K. (2002). Recognition of emotion from facial, prosodic, and written verbal stimuli in Parkinson's disease.*Cortex*, *38*, 623-630.

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.

Katsikitis, M., &Pilowsky, I. (1996).A controlled study of facial mobility treatment in Parkinson's disease.*Journal of Psychosomatic Research, 40*, 387-396.

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.

Keltner, D. (1995). Signs of appeasement: Evidence for the distinct displays of embarrassment, amusement, and shame. *Journal of Personality and Social Psychology, 68,* 441-454.

Kipp, M. (2001) Anvil - A Generic Annotation Tool for Multimodal Dialogue.*Proceedings of the*

7th European Conference on Speech Communication and Technology (Eurospeech), pp. 1367-1370.

Kraut, R. E., & Johnston, R. E. (1979). Social and emotional messages of smiling: An ethological approach. *Journal ofPersonality and Social Psychology, 37,* 1539-1553.

Krumhuber, E., Manstead, A.S.R., &Kappas, A. (2007). Temporal aspects of facial displays in person and expression perception: The effects of smile dynamics, head-tilt, and gender. *J Nonverbal Behav, 31*, 39-56.

Kulisevsky, J., Pagonabarraga, J., Pascual-Sedano, B., Gironell, A., Garcia-Sanchez, C., Martinez-Corral, M. (2008). Motor changes during sertraline treatment in depressed patients with Parkinson's disease. *European Journal of Neurology, 15*, 953-959.

Lafrance, M., & Hecht, M. A. (1999). Option or obligation to smile: The effects of power and gender on facial expression. In P. Philippot, R. S. Feldman, & E. J. Coats (Eds.), *The social context of nonverbal behavior* (pp. 45-70). New York: Cambridge University Press.

LaFrance, M., Hecht, M.A., &Paluck, E.L., (2003). The contingent smile: A meta-analysis of sex differences in smiling. *Psychological Bulletin, 129*, 305-334.

Lebensohn, Z. M. & Jenkins, R. B. (1975). Improvement of Parkinsonism in depressed patients treated with ECT. *TheAmerican Journal of Psychiatry, 132*, 283-285.

LeDoux, J. E. (2000). Emotion circuits in the brain.*Annual Review of Neuroscience, 23*, 155-184.

Liotti, M., Vogel, D., Ramig, L., New, P., Cook, C., & Fox, P. (2003). Functional reorganization of speech-motor function in Parkinson disease following LSVT: A PET study. *Neurology,60*, 432-440.

MacLean, P.D. (1949).Psychosomatic disease and the "visceral brain:" Recent developments bearing on the Papez theory of emotion.*Psychosomatic Medicine, 11,* 338-353.

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

Marsh, L. (2008). Parkinson's disease, In C. G. Lyketsos, P. V. Rabins, & J. R. Lipsey (Eds.), *Psychiatric aspects of neurologic diseases* (pp. 189-216). New York: Oxford University Press.

Matsumoto, D., Keltner, D., Shiota, M. N., O'Sullivan, M., & Frank, M. (2008).Facial Expressions of Emotion.In M. Lewis, J. Haviland-Jones, & L. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 211-234). New York: Oxford University Press.

Maxwell, S.E., &Delaney, H.D. (1990). *Designing experiments and analyzing data: A model comparison perspective.* Belmont, CA: Wadsworth.

Mayberg, H. S., & Solomon, D. H. (1995). Depression in Parkinson's disease: A biochemical and organic viewpoint. *Advances in Neurology, 65*, 49-60.

Mayberg, H. S., Starkstein, S. E., Sadzot, B., Preziosi, T., Anderzejewski, P. L., Dannals, R. F. et al. (1990). Selective hypometabolism in the inferior frontal lobe in depressed patients with PD. *Annals of Neurology, 28,* 57-64.

Mayeux , R., Stern, Y., Rosen, J., &Leventhal J. (1981). Depression, intellectual impairment, and Parkinson disease.*Neurology, 31*, 645-650.

Mayeux, R., Stern, Y., Williams, J. B., Cote, L., Frantz, A., Dyrenfurth, I. (1986). Clinical and biochemical features of depression in Parkinson's disease.*American Journal of Psychiatry, 143*, 756-759.

McCabe, D.L., Borod, J.C., Meltzer, E., Spielman, J., & Ramig, L. (2010). Masked facies in Parkinson's disease: Emotional and motoric factors, neuropathology, Duchenne smiling, and treatment. In A. Freitas-Magalhaes (Ed.), *Emotional expression: The brain and the face,* Vol. 2. Porto, Portugal: University Fernando Pessoa Press.

McCabe, D., Creighton, J., Twaite, J. T., Teague, E. B., Spielman, J., Scorpio, K., . . . Borod, J. C. (2011, October 4). Aprosodia (Update). In S. Gilman (Editor-in-Chief), *MedLink Neurology*. San Diego, CA: MedLink Corporation. Available from MedLink Neurology Website, http://www.medlink.com.

Mergl, R., Mavrogiorgou, P., Hegerl, U., & Juckel, G. (2005). Kinematical analysis of emotionally induced facial expressions: A novel tool to investigate hypomimia in patients suffering from depression. *Journal of Neurology, Neurosurgery, & Psychiatry, 76*, 138-140.

Miller, K.M., Okun, M.S., Marsiske, M., Fennell, E.B., & Bowers, D. (2009).Startle reflex hyporeactivity in Parkinson's disease: An emotion-specific or arousal-modulated deficit? *Neuropsychologia, 47,* 1917-1927.

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.

Morecraft, R.J., Louie, J.L., Herrick, J.L., & Stilwell-Morecraft, K.S. (2001). Cortical innervation of the facial nucleus in the non-human primate: A new interpretation of the effects of stroke and related subtotal brain trauma on the muscles of facial expression. *Brain, 124*, 176-208.

Morecraft, R.J., Stilwell-Morecraft, K.S., &Rossing. (2004). The motor cortex and facial

expression: New insights from neuroscience. *The Neurologist, 10,* 235-249

Narayana, S., Fox, P. T., Zhang, W., Franklin, C., Robin, D., Vogel, D., et al. (2009). Neural correlates of efficacy of voice therapy in Parkinson's disease identified by performance correlation analysis. *Human Brain Mapping, 31,* 222-236.

Olanow, C. W. Stern, M. B. &Sethi, K. (2009). The scientific and clinical basis for the treatment of Parkinson disease.*Neurology, 72*, S1-S136.

Pell, M.D. & Leonard, C.L. (2003). Processing emotional tone from speech in Parkinson's disease: a role for the basal ganglia. *Cognitive Affective and Behavioral Neuroscience*, *3*, 275-278.

Pell, M.D, Cheang, H.S., & Leonard, C.L. (2006). The impact of Parkinson's disease on vocal-prosodic communication from the perspective of listeners. *Brain and Language*, *97*, 123-134.

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.

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.

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.

Provine, R. R., & Fischer, K. R. (1989). Laughing, smiling, and talking: relation to sleeping and social context in humans. *Ethology, 83*, 295-305.

Ramig, L. O., Countryman, S., O'Brien, C., Hoehn, M., & Thompson, L. (1996). Intensive speech treatment for patients with Parkinson's disease: Short and long term comparison of two techniques. *Neurology, 47*, 1496-1504.

Ramig, L.O., Countryman, S., Thompson, L., & Horii, Y. (1995). Comparison of two forms of intensive speech treatment for Parkinson disease. *Journal of Speech and Hearing Research, 38*, 1232-1251.

Ramig L. O., Pawlas, A., & Countryman, S. (1995). *The Lee Silverman voice treatment (LSVT®): A practical guide to treating the voice and speech disorders in Parkinson disease*. Iowa

City, IA: NationalCenter for Voice and Speech.

Ramig, L. O., Sapir, S., Countryman, S., Pawlas, A. A., O'Brien, C., Hoehn, M., & Thompson, L. (2001). Intensive voice treatment (LSVT) for patients with Parkinson's disease: A 2 year follow up. *Journal of Neurology, Neurosurgery, & Psychiatry, 71*, 493- 498.

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

Ravina, B., Edwards, E., & Sheehy, P. (2003). *Summary – diagnosis of depression in Parkinson's disease*. Retrieved August 5, 2009, from National Institute of Neurological Disorders and Stroke Web site: http://www.ninds.nih.gov/news_and_events/proceedings/depression_summary.htm

Reid, S. A. (2000). The experience of emotion in Parkinson's disease and normal aging: Assessing the roles of facial expressiveness and cognitive functioning. Dissertation Abstracts International: Section B. Sciences and Engineering, 60, 4905.

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.

Robert, P. H., Berr, C., Volteau, M., Bertogliati, C., Benoit, M., Sarazin, M., et al. (2006). Apathy in patients with mild cognitive impairment and the risk of developing dementia of Alzheimer's disease.*Clinical Neurology and Neurosurgery, 108*, 733-736.

Root, A. A., & Stephens, A. (2003). Organization of the central control of facial expression in man. *The Journal of Physiology, 549*, 289-298.

Ross, R. &Mathiesen, R. (1998) Volitional and emotional supranuclear facial weakness.*New England Journal of Medicine,* 338, 1515.

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.

Sandyk, R. &Derpapas, K. (1993). Further observations on the unique efficacy of picoTesla range magnetic fields in Parkinson's disease. *International Journal of Neuroscience, 69*, 167-183.

Sano, M., Stern, Y., Cote, Williams, J., L., &Mayeux R. (1990). Depression in Parkinson's disease: A biochemical model. *The Journal of Neuropsychiatry and Clinical Neurosciences, 2*, 88-92.

Sano, M., Stern, Y., Williams, J., Cote, L., Rosenstein, R., &Mayeux R. (1989). Coexisting dementia and depression in Parkinson's disease.*Archives of Neurology, 46*, 1284-1286.

Sapir, S., Ramig, L. O., Hoyt, P., Countryman, S., O'Brian, & Hoehn, M. (2002). Speech loudness and quality 12 months after intensive voice treatment (LSVT) for Parkinson's disease: A comparison with and alternative speech treatment. *Folia PhoniatricaetLogopaedica, 54*, 296-303.

Sapir, S., Spielman, J., Ramig, L. O., Story, B., & Fox, C. (2007). Effects of intensive voice treatment (LSVT) on vowel articulation in dysarthric individuals with idiopathic Parkinson's disease: Acoustic and perceptual findings. *Journal of Speech, Language, and Hearing Research, 50,* 899-912.

Sartorio, E., De Gaspari, D., Baietto, C., Siri, C., Florita, M., Pezzoli, G., et al. (2005). Relationship between Parkinson's disease and neuropsychological symptoms of depression. *Clinical Neuropsychiatry: Journal of Treatment and Evaluation, 2*, 174-179.

Schneider, F., Habel, U., Volkmann, J., Regel., S. Kornischka, J, Sturm, V., & Freund, H. J. (2003). Deep brain stimulation of the subthalamic nucleus enhances emotional processing in Parkinson's disease. *Archives of General Psychiatry, 60*, 296-3002.

Schmidt, K.L., Cohn, J.F., &Tian, Y. (2003). Signal characteristics of spontaneous facial expressions: automatic movement in solitary and social smiles. *Biological Psychiatry, 65*, 49-66.

Schulz, G. M., & Grant, M. K. (2000). Effects of speech therapy and pharmacologic and surgical treatments on voice and speech in Parkinson's disease: A review of the literature. *Journal of Communication Disorders, 33*, 59-88.

Serby, M. (1980). Psychiatric issues in Parkinson's disease.*ComprehensivePsychiatry, 21*, 317-322.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples).*Biometrika 52,* 591–611.

Shulman, L. M., Taback, R. L., Rabenstein, A. A., Weiner W. J. (2002). Non-recognition of depression and other non-motorsymptoms in Parkinson's disease. *Parkinsonism & Related Disorders, 8*, 193-197.

Siemers, E. R., Shekhar A., Quaid K., & Dickson, H. (1993). Anxiety and motor performance in Parkinson's disease.*Movement Disorders, 8*, 501-506.

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.

Simple Boxplot Summaries of Separate Variables (n.d.). In IBM SPSS Statistics. Retrieved from http://publib.boulder.ibm.com/infocenter/spssstat/v20r0m0/index.jsp?topic=%2Fcom.ibm.spss.statistics.help%2Fidh_gxss.htm

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.

Soussignan, R., &Schaal, B. (1996). Children's facial responsiveness to odors: Influences of odor valence, gender, age, and social presence. *Developmental Psychology, 32,* 367-379.

Soussignan, R. (2002). Duchenne smile, emotional experience, and autonomic reactivity: A test of the facial feedback hypothesis. *Emotion, 2*, 52-74.

Spielman J., Borod J. C., Ramig L. (2003). Effects of intensive voice treatment (LSVT) on facial expressiveness in Parkinson's disease: preliminary data. *Cognitive and Behavioral Neurology, 16*, 177-188.

Sprengelmeyer, R., Young, A.W., Mahn, K., Schroeder, U., Woitalla, D., Büttner, T., Kuhn, W., &Przuntek, H. (2003). Facial expression recognition in people with medicated and unmedicated Parkinson's disease. *Neuropsychologia*, *41*, 1047-1057.

Starkstein, S. E., Preziosi, T. S., Bolduc, P. L., & Robinson, R. G. (1990). Depression in Parkinson's disease.*The Journal of Nervous and Mental Disease, 178*, 27-34.

Starkstein, S. E., Mayberg, H. S., Preziosi, T. J., Andrezejewski, P., Leiguarda, R., & Robinson, R. G. (1992). Reliability, validity, and clinical correlates of apathy in Parkinson's disease.*The Journal of Neuropsychiatryand Clinical Neurosciences, 4*, 134 -139.

Stein, M. B., Heuser, I. J., Juncos, J. L., Uhde, T. W. (1990). Anxiety disorders in patients with Parkinson's disease. *American Journal of Psychiatry,147*, 217-220.

Stochl, J., Boomsma, A., Ruzicka, E., Brozova, H., &Blahus, P. (2008). On the structure of motor symptoms of Parkinson's disease.*Movement Disorders, 23*, 1307-1312.

Taylor, A. E., & Saint-Cyr, J. A. (1990). Depression in Parkinson's disease: Reconciling physiological and psychological perspectives. *Journal of Neuropsychiatry, 2*, 92-98.

Tekin, S. & Cummings, J. L. (2002). Frontal-subcortical neuronal circuits and clinical neuropsychiatry: An update. *Journal of Psychosomatic Research, 53,* 647-654.

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.

Treciokas, L. J., Ansel, R. D., & Markham, C. H. (1971). One to two year treatment of

Parkinson's disease with levodopa.*California Medicine, 114*, 7-14.

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

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.

Wooten, G.F., Currie, L.J., Bovbjerg, V.E., Lee, J.K., &Patrie, J. (2004). Are men at greater risk for Parkinson's disease than women? *Journal of Neurology, Neurosurgery, and Psychiatry, 75*, 637-639.

Yecker, S., Borod, J. C., Brozgold, A., Martin, C., Alpert, M., Welkowitz., J. (1999). Lateralization of facial emotional expression in schizophrenic and depressed patients.*The Journal of Neuropsychiatry and Clinical Neurosciences, 11*, 370-379.

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.

Zgaljardic, D. J., Borod, J. C., Foldi, N. S., Mattis, P. (2006). An examination of executive dysfunction associated with frontostriatal circuitry in Parkinson's disease. *Journal of Clinical and Experimental Neuropsychology, 28*, 1127-1144.

Zgaljardic, D. J., Borod, J. C., Foldi, N. S., Rocco, M., Mattis, P. J., Gordon, M. F., Feign, A. S., &Eidelberg, D. (2007). Relationship between self-reported apathy and executive dysfunction in nondemented patients with Parkinson disease.*Cognitive and Behavioral Neurology,* 20, 184-192.

Zgaljardic, D. J., Foldi, N. S., & Borod, J. C. (2004). Cognitive and Behavioral dysfunction in Parkinson's disease: Neurochemical and clinicopathological contributions. *Journal of Neural Transmission, 111*, 1287-1301.