INTRODUCTION
Parkinson’s disease (PD), a progressive neurodegenerative disorder, manifests motor symptoms including rest tremor, bradykinesia, and rigidity [1]. Essential Tremor (ET) is a neurological disorder characterized by postural and kinetic tremor [2]. Clinicians commonly rate ET or PD tremor severity (TS) on the 0 to 4 Fahn-Tolosa-Marin Tremor Rating Scale (TRS) [3] based on observation of tremor, motor ability by tracing Archimedes spirals [4], and self-reported functional disability [5]. A problem with the TRS is subjectivity due to inter-rater and self-reporting variability [6,7]. Quantitative tremor measurement could improve rating reliability because it is objective and context independent. Therefore, the study aim was to capture objective tremor data, quantify TS from the data, and assess the ability of different computational methods to classify TS by matching them to clinical diagnostic assessments.

METHODS
Ten participants were recruited of which three exhibited a history of ET, six PD, and one both PD and ET. Each participant had a Deep Brain Stimulation (DBS) implant and was rated for TS with DBS both on and off by clinicians using TRS and by objective tremor assessment.

Participants wore a miniature 3-axis gyroscope, the Shimmer Wireless Sensor Unit (SWSU), on the wrist while extending arm full length then touching nose five times, holding hand at horizontal extension for five seconds, resting hand in lap for five seconds, and tracing a printed spiral. Gyroscope data were numerically analyzed to determine raw gyroscope data (°/sec), power spectral density (PSD), (°/sec)² / Hz, peak frequency and magnitude, and root mean squared (RMS) value. From this data, TS was calculated on a 0-4 scale (Figure 1) using the SWSU method: determining maximum and minimum RMS values; and machine learning (ML) method: using the classifiers random forest, decision tree, nearest neighbor, multilayer perceptron, and support vector machine. Classifier accuracy in predicting TS was determined using 10-fold cross validation.

From digital pen spiral tracings, five factors were derived characterizing deviations from the printed spiral radius: maximum difference (Δr_max), average difference (Δr_avg), squared Pearson product moment correlation coefficient (R²), RMS, and standard deviation (σ). Equations calculating averaged TS from the 5 factors were derived by plotting the 5 factors against TS ratings (digital pen method).

RESULTS AND DISCUSSION
We compared the SWSU and spiral TS ratings to the clinicians’ TS ratings. Clinician ratings matched SWSU ratings with 42%, digital pen 78%, and ML (decision tree algorithm) 82% accuracy. It is important to note that a value equal to 20% is due to random effect. These results represent a substantial improvement over the clinical inter-rater reliability for the TRS (Kappa statistic =~0.5) [5]. Because the ML algorithms used the clinical rating from three different raters as the target data, a match of 100% would not be expected due to inter-rater variability. Although the upper limit for matching clinical ratings was not established due to the limited number of subjects, the 82% match (ML algorithms) and 78% match (digital pen) shows that these methods have the capability to provide more reliable assessments of the TRS scale.

CONCLUSIONS
TS was objectively rated with greatest reliably using a ML decision tree classifier with RMS tremor gyroscope data. This computational method has the potential to substantially increase the reliability of tremor assessment.

REFERENCES

ACKNOWLEDGEMENTS
Support by: Adapx Inc. and NSF Grant No. DGE-090078