Executive Function Training to Reduce Cognitive Intra-Individual Variability in Adults with HIV

R21AG077957

NCT05598047

The EFT Study
(Executive Functioning Training)

January 3, 2022

David E. Vance, PhD, MGS, FAAN

RESEARCH PLAN

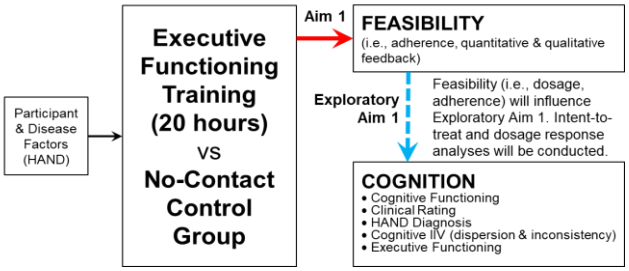
A. SIGNIFICANCE

A1. Importance of Cognition

Addressing HAND among PWH is vital. **1)** HIV accentuates/accelerates aging,[55,56] including cognitive aging; and these impairments occur across cognitive domains although executive functioning is particularly affected.[13,27-40,57-59] In a 2018 meta-analysis of 37 studies of PWH, "executive functioning may be differentially affected by HIV" and should be specifically targeted for intervention.[45] **2)** Executive functioning mediates the detrimental effects of white matter hyperintensities and other brain insults for all affected domains.[46,60] **3)** Cognitive impairments in HIV are related to poorer everyday functioning and quality of life.[70,80,61] **4)** Few behavioral interventions aimed at improving cognition in this pharmacologically-burdened population have been attempted.[62] **5)** Computerized cognitive training interventions have been shown to improve cognition across populations.[50,63,64] Improvement in individual cognitive domains such as executive functioning resulting from cognitive training translate into improved performance on everyday functioning and quality of life.[65-68]

Figure 1. Specific Aims and Conceptual Model

Specific Aim 1: Red/Solid
Exploratory Aim 1: Blue/Dashed



A2. Cognitive IIV in HIV & Computerized Cognitive Training

Fluctuations in an individual's cognitive ability across cognitive domains (dispersion) or across multiple trials within the same test (inconsistency) may be more reflective of underlying pathology than traditional neuropsychological measures. These fluctuations are known as **Intra-Individual Variability (IIV)**. In our systematic review of 13 cognitive IIV studies in PWH, greater cognitive IIV predicted cognitive decline as well as decreased cognitive integrity and neuropathology.[43,44,47] For example, in 126 PWH and 40 HIV-negative adults, Morgan et al.[44] found significant HIV status X age group interactions, with the older HIV+ group showing the greatest IIV (dispersion). The **Executive Dyscontrol Hypothesis** posits that cognitive IIV may result from "executive dyscontrol"[43] or the inefficiency with which executive control processes are unable to coordinate other cognitive processes/domains.[45,46] IIV increases with normal aging and greater IIV is associated with increasing cognitive impairment and characteristic changes in neuroimaging studies.[42,53]

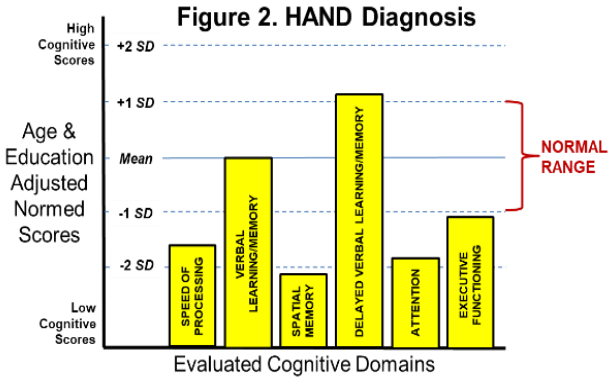
Cognitive training also produces changes in brain function and morphology, resulting in cognitive improvement.[63,64,69,70] Over a 5-8 week period, Chang et al. administered ~10 hours of working memory training (a type of executive functioning training) to 173 PWH and seronegative participants.[71] Although the PWH had poorer cognitive functioning compared to the seronegative group at baseline, those who received the working memory training improved. Compared to the seronegative group at baseline, fMRI revealed PWH experienced greater frontal activation while performing tasks of working memory, suggesting that greater neural effort was needed to compensate for poorer cognitive functioning. Yet, both the seronegative group and PWH had decreased frontal brain activation 1 to 6 months after training, indicative of improved neural efficiency. Unfortunately, cognitive IIV was not examined, nor was 20 hours of training administered which would likely have resulted in more robust findings. Yet, these results suggest that executive functioning training may reduce strain on executive control resources; then those neural resources can be reallocated to other cognitive domains needing compensation, as a result such training should reduce cognitive IIV and improve cognition.

B. INNOVATIONS

This proposal translates noninvasive, inexpensive, accessible, computerized cognitive training exercises to improve cognition and reduce the prevalence/severity of HAND. Furthermore, this proposal is innovative in including cognitive IIV as the guide to target executive functioning. Cognitive IIV will also serve as an outcome of treatment, which has never been attempted before in the neuroHIV literature.[47,49]

B1. Framework for Targeting the Cognitive Training & Diagnosing HAND

Based on a test battery of at least 6 normed cognitive domains (see Tables 1 & 4) administered to the participant, the level of HAND is determined ranging from least severe (Asymptomatic Neurocognitive Impairment, ANI) to most severe (HIV-Associated Dementia, HAD), with Mild Neurocognitive Disorder (MND) falling in between. ANI and MND, the most common levels of HAND (33% & 12%, respectively), are diagnosed when a person scores more than 1 *SD* below her adjusted norm in two cognitive domains;[3,4] in Figure 2, this is observed for speed of processing, spatial memory, attention, and executive functioning. Although rare (<2%), HAD occurs when a person performs 2 or more *SD* below her adjusted norm in at least two cognitive domains. In Figure 2, this person performed more than 2 *SD below her adjusted norm in only one domain,* spatial memory, and thus does not meet the criteria for HAD.[3,4] Yet, if she performed slightly worse on speed of



processing or attention such that her score was more than 2 *SD below the adjusted norm*, then a diagnosis of HAD would be assigned.

In the case of an individual with impairments in three or more domains, it is not feasible to apply 10 hrs of training to each domain; this would represent an unreasonable burden of cost, time, and patient fatigue. Hence, it behooves clinicians and researchers to be selective in the domains targeted for training. Many studies employ between 10 to 20 hrs of cognitive training; and additional hrs of training often result in diminished cognitive gains and participant fatigue.[69] A practical strategy is to limit training to 20 hrs maximum).

In this example (Figure 2), speed of processing, spatial memory, attention, and executive functioning all contribute to a diagnosis of HAND. So which cognitive domains should be targeted for training? Even with 10 hours of training in each impaired cognitive domain, this represents 40 hours of training which is very demanding on the individual. Also, coordinating 4 different types of cognitive training can result in administration error and be confusing to participants. **Fortunately, both the Executive Dyscontrol Hypothesis and Cognitive IIV point to which cognitive domain to target for intervention**.[43,46]

Solution: Executive Functioning Training – Improving executive functioning ability/resources can improve functioning in other cognitive domains. In our prior work, executive functioning training not only improved executive functioning (*n*=9; *d*=-0.89), improvements were also observed in speed of processing (*d*=-0.56), attention (*d*=-1.24), and delayed verbal learning and memory (*d*=-0.40).[72] **Interpretation –** Executive functioning training improved the coordination of other cognitive domains.

B2. Dispersion vs Inconsistency

Dispersion and inconsistency are considered equivalent forms of cognitive IIV, but actually there is little in the literature comparing the predictive validity of these two types. Methodologically, we will be uniquely equipped to do this comparison by including both dispersion and inconsistency (Table 4) in the same study, making this an unprecedented contribution to the cognitive IIV field.[49]

B3. Previous Work/Pilot Studies

Partnering with the UAB Center for AIDS Research, our work has resulted in 100+ articles.[61,66,67,73-79]

B3.1. Systematic Review on IIV and HIV[47] – In our systematic review of 13 studies on cognitive IIV in PWH, IIV was related to poorer global cognitive functioning and often *poorer executive functioning*.

B3.2. Systematic Review of Cognitive Training in HIV[49] – In our systematic review of 13 cognitive training studies in PWH, cognitive training improved the domain that was targeted (e.g., executive functioning training improved executive functioning). Yet, none of these training studies incorporated IIV into either the design of the training protocol or as an outcome measure. Our R21 proposal is significant and innovative in that we use IIV to target our intervention (i.e., treating executive dysfunction) and consider it as an outcome.

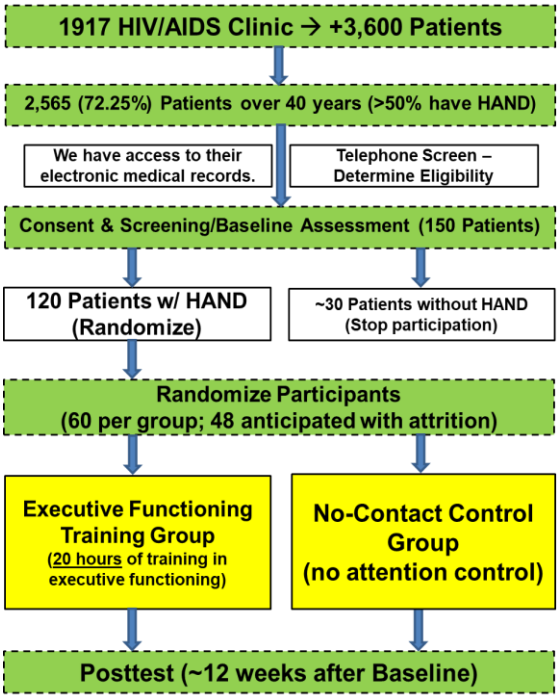
B3.3. Training On Purpose Study[48] – In our study, 88 PWH with HAND were randomized to either: 1) a no-contact control group (*n*=40) or 2) the Individualized-Targeted Cognitive Training group (*n*=48). Based on Salthouse's[80,81] and Wicken's[82] theories of information processing, our prior theoretical framework was used to determine which two cognitive domains were selected for each participant's training. Participants received speed of processing training and/or attention training if their profile included any degree of impairment in these domains; if not, participants were assigned to cognitive training in the one/two **least** impaired cognitive domains that contributed to their HAND diagnosis. Unexpectedly, there was no group effect of training on the global cognitive score (*p*=0.256; *d*=-0.21). However, when examining individualized cognitive training and respective cognitive domains, improvements were observed following executive function training (*n*=9; *d*=-0.89). **Upon Further Reflection –** Based on the cognitive IIV literature and Executive Dysfunction Hypothesis,[46,83,84] our new approach may be more therapeutically beneficial by solely targeting executive functioning training, rather than the least impaired cognitive domains. By training executive functioning, we expect to reduce cognitive IIV (indicating improvement in executive coordination), which may then improve global cognition and decrease the prevalence and severity of HAND. Our shift in theory and methodology makes this approach highly likely to be effective – *making this high-risk high-reward R21 application promising.*

C. APPROACH

C1. Design Overview

A pre-post three-group experimental design will be used (Figure 3). Participants will be recruited from the UAB 1917 HIV/AIDS Clinic which has a patient population of +3,600 and is the largest HIV medical provider within 100 miles. Eligible participants (see Section C2) will be consented at the UAB Center for Research on Applied Gerontology where a ~2 hr baseline assessment will be administered. Participants' neuropsychological data gathered at baseline will be examined to determine a HAND classification. Only participants with HAND (see Section B1) will be invited to continue with the study. Stratified random assignment will ensure an equal number of participants in each group by gender, minority status, and with/without executive functioning impairment (i.e., 1 *SD* below normative mean). After training, participants will complete a posttest assessment.

Figure 3. Study Design



C2. Recruitment and Retention Procedures

Inclusion Criteria. Participants (men & women) must be 40+ years, English speaking, and have HAND. PWH 40 years of age and older are targeted due to the greater HAND prevalence with middle and older age.[41]

Exclusion Criteria. Participation requires ~12 weeks and in-person visits; participants living beyond 60 miles away from the center will be excluded. Participants living in unstable housing (e.g., halfway house) or with significant neuromedical comorbidities (e.g., schizophrenia) will be excluded. Other conditions (e.g., legally blind/deaf, currently undergoing radiation or chemotherapy, or a history of significant brain trauma, diagnosed with COVID-19 over the past 3 months[85]), that could impact cognitive functioning or testing also necessitate exclusion. These criteria are typical of neuroHIV studies.[5,13,86-89] As cognitive training effects can be robust two years after training is completed, those who have received cognitive training within the past three years will be excluded.

HAND Criteria. We are targeting those with HAND as they will have room to improve cognitively. We rarely enroll those with HIV-Associated Dementia as they tend to self-exclude themselves and is rare (<2%).[48]

Exceptions and Ecological Validity. From electronic medical data from this same clinic,[62] many older adults with HIV experience depression (~45%), anxiety (~20%), substance abuse (~24%), alcohol abuse (~15%), and hepatitis C (~20%), all of which can impact cognition and cognitive training.[5,90-93] It is not feasible to exclude PWH with such conditions; likewise, including such participants will increase ecological validity and generalizability of the findings, which has been a strength of our prior studies.[26,48,78] Our randomization procedures control for such potential confounds; if needed, ad hoc analyses can provide additional insights into the impact of these covariates on adherence/feasibility and treatment effects.

Recruitment/Retention Strategies. As effective in our studies,[26,48,94-101] recruitment/retention strategies will be used. **1)** Recruitment materials distributed in the UAB 1917 HIV/AIDS Clinic. **2)** Compensation will be provided ($50 for baseline assessment, $15 for every hr of training, & $75 for posttest assessment). **3)** Reminder letters and telephone calls will be used before the scheduled visit. **4)** Beverages/snacks will be provided (from departmental funds). **5)** Secondary contact information will be collected to follow up with lost participants.

C3. Intervention Protocol

BrainHQ (POSIT Science Inc.) cognitive training modules will be used as in our other studies (Table 1); these programs have gaming components that encourage adherence. BrainHQ cognitive training products are tested and endorsed by the scientific community (see letter of support). A meta-analysis[50] of computerized cognitive training in older adults found optimal therapeutic effects occurred when training sessions last at most 60 minutes and are administered 1-3 times per week – dosage parameters already incorporated in our study. This self-administered program uses touch-screen technology with tablets which allows computer novices to engage with the training exercises. The intervention will be administered in the research lab of the UAB Center for Applied Gerontology where many of our other studies have been conducted. Working with BrainHQ, when logging on, participants can only receive the individualized cognitive training exercises they are assigned.

C3.1. Executive Functioning Training Group.

As seen in Table 1, those in the Executive Functioning Training Group will engage in exercises requiring one to set shift; that is, to maintain at least two sets of rules and decide which is appropriate to determine the response. In using these training exercises in the TOPS study, the effects size was quite large (*d*=-0.89).

Dosage of 20 hours of training is considered an upper range on how much training is needed to produce an optimal therapeutic effect.[48-50]

Table 1. Executive Functioning Training Exercises

Cognitive Training Domain	Cognitive Training Exercises Matching the Cognitive Domain (see POSIT Science Inc (Brain HQ) at www.brainhq.com for details)
Executive Functioning – These exercises require one to set shift; that is, to maintain at least two sets of rules and decide which one is appropriate for the response. The effect size for executive functioning training in our last study was *d* = -0.89.[48]	**Mind Binder –** This exercise is a "set shifting" exercise where one is presented with two rules, but one must choose the correct answer based upon the rules provided in this exercise. **Mixed Signals –** This exercise requires one to listen to a number, letter, or other information while looking at a similar set of information, similar to the Stroop test. **Card Shark –** This exercise is an extension of a visual n-back paradigm using an aspect of executive functioning (i.e., working memory). **Freeze Frame –** This exercise is an extension of the go/no go paradigm using an aspect of executive functioning (i.e., working memory).

C3.2. No-Contact Control Group. This group will receive no intervention. As this is a pilot/feasibility study, we do not have the resources to provide a contact control group. Importantly, in a prior study, we established that a no-contact control group and a contact control (sham) group did not significantly differ from each other and both served as an excellent comparison to a cognitive intervention.[79]

C3.3. Treatment Fidelity. A treatment fidelity checklist will be used to review the amount of time engaged in training with participants. Also, the BrainHQ (i.e., POSIT Science) software monitors the time participants spend engaged in each exercise. In the seminal ACTIVE Study (*N*=2,802, the largest clinical trial of cognitive training), participants are considered "trained" when they successfully complete 80% of training.[75] Like our other studies,[102] both a completers-only and an intent-to-treat analysis will be conducted. The completers-only analysis is appropriate for use when examining the actual potential of the cognitive training. If participants do not complete training, their data will still be examined using an intent-to-treat analysis.

C4. Instruments (Tables 2 - 4; Administered Baseline and Posttest)

Administration time of the assessment will be ~2 hrs. We will use REDCap and BRACE+ tablet for administration of the instruments below to reduce tester burden, tester error, and improve the efficiency of data entry and data management, which will save significant staff time and resources. Other cognitive studies as observed in the HIV Neurobehavioral Research Center (HNRC) group employ testing assessments of similar length.[103-105] BRACE+ was develop in conjunction with HNRC and Dr. Leah Rubin (see letter of support).

The assessment battery is based on the conceptual model in Figure 1. Because participants will be recruited through the UAB 1917 (HIV/AIDS) Clinic, the following information will be available for analysis from the electronic medical records through the CFAR Statistical Core. Testing will be reviewed and observed every quarter to prevent drift in the protocol.

Table 2. Demographic, Background, and Covariate Measures (*Baseline Only)
***Demographic Questionnaire –** Used to gather basic background characteristics (e.g., gender, ethnicity, etc). (2 min.)
***Wide Range Achievement Test-3 (WRAT-3) –** Used to determine educational quality.[106] (2 min.)
Addiction Severity Index (ASI) – A widely used, gold standard measure of alcohol and drug use.[107,108] (4 min.)
***HIV History/Status –** Used to assess current HIV status (i.e., viral load) and general HIV infection history. (2 min.)
Centers for Epidemiological Studies – Depression (CES-D) scale measures how often participants acknowledge 20 verbal depressive symptoms. Cronbach's α is very good at 0.88.[109] (2 min.)
Electronic Medical Records (EMR) – EMR is accessed through the 1917 HIV/AIDS Clinic as in our prior studies. Such data include: viral load, CD4+ count, lipid profile, etc. Lab values that correspond most closely to the study visit will be used (i.e., within 90 days).

C4.1. Feasibility and Acceptability Measures

Table 3. Aim 1 Measures – FEASIBILITY & ACCEPTABILITY
Baseline Assessment of Cognitive Training – All participants at baseline will be asked questions about computer use, knowledge about cognitive training, their perceptions about whether they need cognitive training, knowledge about HIV and cognition, etc. Both quantitative and qualitative (i.e., open-ended responses) data will be collected; this is similar to our other studies.[98,110-112] (10 min.)
Cognitive Training Satisfaction (after 10 hours of training completed & posttest only) – Used to assess likes/dislikes of the intervention, both quantitative questions and qualitative (i.e., open-ended responses) data are gathered as has been used in our previous cognitive intervention studies to evaluate feasibility and acceptability.[111] This is being administered half way through the training for those in the intervention group to safeguard should they withdraw from the study prematurely; this information could prove useful in establishing why they withdraw. (10 min.)
Exit Survey – If a participant withdraws, we will administer a brief quantitative and qualitative (i.e., open-ended responses) survey to assess what he/she liked/disliked about the training, how to improve it, and why he/she is withdrawing from the study (5 min.)
Attrition & Adherence Rates (after study data collection) – We will calculate the attrition rate and adherence rate of the protocol, similar to our other studies.[72,97,111]

C4.2. Cognitive Measures. Both mean-level cognitive functioning (e.g., domain specific and global cognitive functioning) and cognitive IIV calculations will be used to examine cognitive change.

Table 4. Exploratory Aim 1 Measures – COGNITION

BRACE+ (BrainBaseline Assessment of Cognition and Everyday Functioning), supported by NIMH R42099964 and Digital Artefacts/UCSD) is a HIPAA compliant tablet-based cognitive assessment platform. This self-administered tool is not literacy dependent (i.e., automated audio/video instructions) and uses validated cognitive tests sensitive to mild-to-moderate cognitive impairments and have demonstrated good validity relative to a comprehensive standard neuropsychological test battery and test-retest reliability. During its development, *T*-scores (mean=50, *SD*=10) were generated using a normative-based regression approach (adjusted for age, sex, race/ethnicity, and education) from the Women's Interagency HIV Study (WIHS, all women) and the Multicenter AIDS Cohort Study (MACS, all men). **The BRACE+ also demonstrated good discriminant validity when differentiating between PWH with and without HAND (using a *T*-score cutoff ≤ 40) on the gold standard cognitive tests**. All tests are designed for repeat administration, with trials from all experimental conditions randomized within each test administration. This leads to an unlimited number of 'versions' of each test and ensures that no two test administrations are alike, limiting stimulus-specific practice effects that plague many current cognitive assessments. **Automatic scoring allows real time detection of HAND impairment.** (50 min.)

Cognitive Domain Composites – In the case of two or more measures being available in a domain (domains below in **BLUE**), a composite score will be created, thus creating a more stable estimate of that cognitive domain which is a standard practice.[113,114]

Executive Functioning Tests	•**Stroop Interference**[115,116] •N-back[116] •Trails B[115-117,116]
Attention Tests	•**Stroop Color**[116] •Flanker[116] •Visual Search[116] •Posner Cueing Task[116]
Speed of Processing Tests	•**Digit Symbol Substitution Test**[116] •Trails A[115-119]
Immediate & Delayed Verbal Learning/Memory Test	•**Verbal Learning & Recall**[116]
Spatial Memory Test	•**Spatial Working Memory**[116]

HAND & Cognitive Intra-Individual Variability (IIV) Measures

HAND & Domain Composite Scores – HAND is diagnosed using a neuropsychological algorithm called the Frascati criteria.[120] Using each participant's test scores combined over domains, an algorithm is used to calculate a Global Clinical Rating score ranging from 1 (*above average*) to 9 (*severe impairment*). Scores from 5-9 in two cognitive domains indicate HAND. Using this continuum, the Global Clinical Rating score of 5 or higher indicates HAND.[113,114]

Cognitive IIV – Recent research posits that cognitive impairments in PWH are not fully captured by traditional mean-based neuropsychological assessment; in fact, fluctuations in one's individual cognitive variability across cognitive domains (**IIV Dispersion**) or within the same test (**IIV Inconsistency**) may be more reflective of underlying pathology, a concept known as cognitive IIV. Greater cognitive IIV predicts cognitive decline as well as decreased cognitive integrity and increased neuropathology.[43,44,47]

IIV Dispersion – Dispersion (IIV_D) is one type of IIV that refers to variability in performance across multiple measures, such as across tests in the neuropsychological test battery (above). There are two accepted formulas: 1) the intra-individual standard deviation (*iSD*), and 2) the coefficient of variation (*CoV*) (Vance et al., in press).[47,121] Both will be calculated at baseline and posttest as an outcome variable; it is hypothesized that IIV Dispersion will decrease in those in the training groups.
Dispersion Calculation – To keep the dispersion calculation from being influenced by more than 1 cognitive domain, only one representative cognitive test (above in **RED**) from each of the domains will be used to prevent any cognitive domain bias.[47,121]

IIV Inconsistency using the Connors' Continuous Performance Test (10 min.) – Inconsistency (IIV_C) is a type of IIV that refers to variability related to a single person's performance on a single task (i.e., within task) across multiple instances or trials, such as the variability observed across a person's distribution of reaction times (RT) on a cognitive task. There are two accepted formulas: 1) the intra-individual standard deviation (*iSD*), and 2) the coefficient of variation (*CoV*) (Vance et al., in press).[47,121] These also will be calculated at baseline and posttest as an outcome variable as a measure of overall improved cognitive function. The Connor's Continuous Performance Test, Second Edition (CPT-II) is a widely accepted test of sustained/selective attention and impulsivity. The CPT-II is the most commonly used test in the cognitive IIV literature to generate the IIV inconsistency coefficients.[122,123]

C5. Sample Size Considerations

The sample size for this study is based on three considerations: **1)** the exploratory rather than confirmatory nature of the project which focuses on feasibility and effect size estimation rather than null hypothesis significance testing (as there have not been cognitive training studies in IIV on HIV); **2)** the effect size obtained in our Training on Purpose Study (TOPS) on PWH for global cognitive score of |d|=0.21 (section B2) congruent with that reported in the meta-analysis by Lampit et al.[50] of computerized cognitive training (g = 0.22) on elderly adults; and **3)** an expected attrition rate of 19.2% based on TOPS, which is congruent with other studies.[77,124] An effective sample size of 48 per group (after attrition; n per group = 60 enrolled) is sufficient for estimating reasonable upper and lower bounds for an effect size, assuming random sampling from the target population.[125] We conducted 5,000 computer simulations of an experiment with a true effect size of 0.22 comparing two groups with n = 48, intra-subject correlation of 0.5, and data distributed following a Student's t (df = 48) distribution. Tabulation of the 16th and 84th percentiles (i.e., 68% of the distribution centered at the mean) of the 5,000 observed effect sizes resulted in the interval 0.02 – 0.42, which provided reasonable lower and upper bounds for the expected effect size in a larger sample. If formal testing is conducted at a traditional level of 0.05, and again assuming intra-subject correlation of 0.5, a sample of 48 per group provides 80% power to detect an effect of 0.57 (a medium effect size).

C6. Data Analysis

C6.1. Data Management and Tracking. The research assistant/study coordinator will double-key enter de-identified questionnaire responses into an encrypted, password protected (front and back end) electronic database called REDCap (**R**esearch **E**lectronic **D**ata **CAP**ture). REDCap is software for building and managing questionnaires and facilitating electronic data collection and storage. It supports a HIPAA best practice, secure web-based application. This database will be accessed on a secure network server that is password protected.

C6.2. Analysis Plan. De-identified data will be extracted from REDCap and analyzed using SPSS, R, and SAS for longitudinal analysis using mixed-effects modeling.[126] Distributional properties of the individual outcomes will be examined, and if necessary methods appropriate for outcomes with non-normal distributions will be employed, such as using transformations or generalized mixed effects modeling. Measures of effect size will be used to assess balance at baseline between the study groups in terms of the study outcomes, demographic variables, depression, substance use, medications, and medical condition variables. Subsequent analyses will use covariate adjustment if relevant imbalance is present among any participant characteristics at baseline. Despite efforts to collect data as complete as possible, missing data due to dropout are expected. The impact of the missing data will be mitigated by using mixed effects modeling with covariate adjustment for baseline characteristics, if any, that are relevantly associated to dropout. In the presence of missing data, mixed effects models provide unbiased estimates of model parameters, as long as the data are missing at random or conditionally at random (i.e., conditional on adjustment for covariates associated to missingness). If necessary, multiple imputation techniques will be used as sensitivity analyses.[127]

Aim 1 – Feasibility: To determine feasibility and acceptability of the intervention (i.e., attrition, feedback).

We will tabulate adherence rate of the protocols and attrition rate. We will use measures of bivariate association to explore for predictors of attrition and adherence. We will tabulate training satisfaction responses and apply qualitative techniques to examine open-ended responses to feedback about training. We will have multiple sources of data (i.e., see Table 3) in which we can triangulate to examine barriers and strengths to the acceptability and feasibility of this intervention in this clinical population.

Exploratory Aim 1 – Cognition: Compare adults who receive Executive Functioning Training to those who receive no training to determine whether they improve in global cognitive ability and overall cognitive IIV.

HAND diagnosis will be assessed post-intervention to determine the proportions of participants who no longer meet HAND criteria. Between-group difference in proportions will be converted into effect size using the method by Chinn[128], and confidence intervals will be estimated. For each neurocognitive outcome measure (Global Score, IIV Dispersion, IIV Inconsistency), longitudinal models will be fitted with a group assignment indicator, time-point indicator, and group by time-point interaction. A random effect for participant will be fitted to account for non-independence among the repeated measures on the same individuals. Linear contrasts will be used to estimate and compare the groups' mean outcome change. The between-group difference in change will be transformed into an effect size with respective confidence interval. If necessary, the models will be adjusted for baseline covariates imbalanced at baseline or associated with attrition. Interpretation will be based on the magnitude of the estimated effect sizes and the resulting width of the respective confidence intervals.

Pitfalls and Alternatives: We recognize several pitfalls. **1)** Regression to the Mean – Although unavoidable, randomization and the use of the control group will provide some protection. **2)** Cognitive Reserve – Those with more cognitive reserve may benefit more from the training; we can examine that in ad hoc analyses. We often examine baseline cognitive level as a predictor of treatment effects. **3)** Unforeseen Covariates – Ad hoc analyses can provide insights of the effects covariates exert on adherence/feasibility and treatment effects.

C7. Rigor, Transparency, and Reproducibility. Background for the scientific premise of this project and preliminary data are described above. We have used rigorous biostatistical methods to ensure that the proposed study is well designed, includes a reasonable sample size congruent with its exploratory purpose, and incorporates randomization and blinding of data collection to reduce bias. In the event of unanticipated issues the team will employ well-thought-out robust and unbiased methodology if changes are necessary to achieve accurate and reproducible analyses, interpretation, and reporting of results. We will recruit participants diverse in gender and race to increase robustness, unbiasedness, and usability results. We will conduct rigorous data entry and validation, and ensure secured data storage and transfer as the project progresses. To ensure reproducibility, after completion of data collection and tabulation of preliminary results, a co-investigator with expertise in quantitative methods (Dr. Wheeler, UAB SON) will conduct an independent analytical review with cleaned de-identified data to verify and validate the results. Pertinent data and study details will be reported in ClinicalTrials.gov and relevant publications. De-identified cleaned data will be shared upon request under the IRB regulation so that others can reproduce the results.

C8. Future Research. The important first step is to validate that executive functioning training has clinical significance in ameliorating HAND and reducing cognitive IIV. After this proof of concept, other studies should examine: 1) underlying changes in neurological structures/pathways using fMRI, and 2) to what degree this approach to cognitive training in PWH improves everyday functioning and quality of life. Future studies may also focus on improving specific cognitive domains other than executive function to determine if the strain on executive functioning is reduced and HAND symptoms alleviated. We considered this approach initially but reasoned that the focus on reducing cognitive IIV by improving executive functioning is a necessary first step.