

A Machine-Learning Based Approach for Predicting Older Adults?
Adherence to Technology-Based Cognitive Training

Zhe He et al.

Deposited 2023-09-27

Citation of published version:

He, Z., Tian, S., Singh, A., Chakraborty, S., Zhang, S., Lustria, M. L. A., Charness, N., Roque, N. A., Harrell, E. R., & Boot, W. R. (2022). A Machine-Learning Based Approach for Predicting Older Adults' Adherence to Technology-Based Cognitive Training. In *Information Processing & Management* (Vol. 59, Issue 5, p. 103034). Elsevier BV. <https://doi.org/10.1016/j.ipm.2022.103034>



Published in final edited form as:

Inf Process Manag. 2022 September ; 59(5): . doi:10.1016/j.ipm.2022.103034.

A Machine-Learning Based Approach for Predicting Older Adults' Adherence to Technology-Based Cognitive Training

Zhe He^{1,2,*}, Shubo Tian^{3,*}, Ankita Singh⁴, Shayok Chakraborty⁴, Shenghao Zhang⁵, Mia Liza A. Lustria^{1,2}, Neil Charness⁵, Nelson A. Roque⁶, Erin R. Harrell⁷, Walter R. Boot⁵

¹School of Information, Florida State University, Tallahassee, Florida USA

²College of Medicine, Florida State University, Tallahassee, Florida USA

³Department of Statistics, Florida State University, Tallahassee, Florida USA

⁴Department of Computer Science, Florida State University, Tallahassee, Florida USA

⁵Department of Psychology, Florida State University, Tallahassee, Florida USA

⁶Department of Psychology, University of Central Florida, Orlando, Florida USA

⁷Department of Psychology, The University of Alabama, Tuscaloosa, Alabama USA

Abstract

Adequate adherence is a necessary condition for success with any intervention, including for computerized cognitive training designed to mitigate age-related cognitive decline. Tailored prompting systems offer promise for promoting adherence and facilitating intervention success. However, developing adherence support systems capable of just-in-time adaptive reminders requires understanding the factors that predict adherence, particularly an imminent adherence lapse. In this study we built machine learning models to predict participants' adherence at different levels (overall and weekly) using data collected from a previous cognitive training intervention. We then built machine learning models to predict adherence using a variety of baseline measures (demographic, attitudinal, and cognitive ability variables), as well as deep learning models to predict the next week's adherence using variables derived from training interactions in the previous week. Logistic regression models with selected baseline variables were able to predict overall adherence with moderate accuracy (AUROC: 0.71), while some recurrent neural network models were able to predict weekly adherence with high accuracy (AUROC: 0.84–0.86) based on daily interactions. Analysis of the post hoc explanation of machine learning models revealed

Corresponding authors: Zhe He, Ph.D., School of Information, College of Communication and Information, Florida State University, Tallahassee, Florida 32306, Tel: 850-644-5775, zhe@fsu.edu.

*Zhe He and Shubo Tian are equal-contribution first-authors

Author Contributions

Zhe He: Conceptualization, Methodology, Data curation, Formal analysis, Validation, Supervision, Project administration, Roles/Writing - original draft, Writing - review & editing; **Shubo Tian:** Conceptualization, Methodology, Data curation, Software, Formal analysis, Validation, Visualization, Roles/Writing - original draft; **Ankita Singh:** Conceptualization, Writing - review & editing; **Shayok Chakraborty:** Conceptualization, Writing - review & editing; **Shenghao Zhang:** Writing - review & editing; **Mia Liza A. Lustria:** Writing - review & editing; **Neil Charness:** Conceptualization, Methodology, Writing - review & editing; **Nelson A. Roque:** Data curation; **Erin R. Harrell:** Data curation; **Walter R. Boot:** Conceptualization, Methodology, Funding acquisition, Project administration, Writing - review & editing.

Competing Interests

None

that general self-efficacy, objective memory measures, and technology self-efficacy were most predictive of participants' overall adherence, while time of training, sessions played, and game outcomes were predictive of the next week's adherence. Machine-learning based approaches revealed that both individual difference characteristics and previous intervention interactions provide useful information for predicting adherence, and these insights can provide initial clues as to who to target with adherence support strategies and when to provide support. This information will inform the development of a technology-based, just-in-time adherence support systems.

Keywords

Cognitive training; Machine learning; Adherence prediction; Just-in-time intervention

1. Introduction

Population aging, in combination with age-related changes in cognition, represents a significant societal challenge. There were about 54 million older adults (aged 65+) living in the United States in 2019 (Administration on Aging, 2020). This number is projected to increase to 81 million by 2040 and 95 million by 2060. The "oldest old" (85+) age group is the fastest growing segment within this rapidly growing older population. Most aging adults can expect to experience cognitive decline (Salthouse, 2010) and even normative cognitive changes can impact the performance of important everyday tasks (e.g., Allaire & Marsiske, 1999; Diehl et al., 1995). However, up to a quarter of older adults will experience mild cognitive impairment (MCI), cognitive decline greater than typical for one's age, while about 10% will develop Alzheimer's disease or related dementias (Langa et al., 2017; Roberts & Knopman, 2013). These pathological changes pose a serious threat to an individual's ability to live independently. In addition to impacts on health and wellbeing, cognitive decline is associated with substantial financial burden. In 2020 in the United States, the estimated cost of care for individuals with Alzheimer's disease was \$305 billion, an underestimate that does not include the indirect costs of care (e.g., informal caregiving; Foldes et al., 2018). Addressing normative, and especially non-normative, changes to cognition is a vitally important goal that has substantial implications for the health, wellbeing, and independence of older individuals, but also has larger societal implications associated with reducing the resources needed to support large numbers of individuals experiencing significant cognitive impairment.

Early detection and intervention are likely key to combatting the impact of age-related cognitive decline. Physical fitness (Kramer & Colcombe, 2018), pharmacological (Rafii & Aisen, 2015), and cognitive training interventions (Rebok et al., 2014) are promising approaches with respect to achieving this aim. Cognitive training, one of the most popular forms of intervention, typically involves computer exercises designed to improve core cognitive abilities important to the performance of everyday tasks. Although findings in the cognitive training literature have generally been mixed, at least some studies suggest promise with respect to the ability of such interventions to enhance cognition and postpone dementia onset (Bahar-Fuchs et al., 2013; Basak et al., 2020; Chiu et al., 2017; Ge et al., 2018; Gray et al., 2021; Kallio et al., 2017; Nguyen et al., 2019; Sala et al., 2019; Simons

et al., 2016; Turunen et al., 2019; Zhang et al., 2019). Unfortunately, adherence to such programs, even in the context of clinical trials testing their efficacy, can be quite low (e.g., Boot et al., 2013; Corbett et al., 2015; Turunen et al., 2019). As such, uncovering effective strategies to promote adherence could have significant implications for the understanding of the efficacy of such interventions in clinical trials as well as for optimizing cognitive training outcomes should cognitive training interventions be discovered to be effective (Rivera-Torres et al., 2019; Sr et al., 2020). If computerized cognitive training ultimately is discovered to be an ineffective method to boost cognition later in life, lessons learned for how to boost adherence to these interventions might provide generalizable insights into how to boost adherence to other technology-based interventions.

The World Health Organization (WHO) defines adherence as “the extent to which a person’s behavior agrees with the recommendations of a health care provider.” It is generally recognized that adherence is influenced simultaneously by a host of individual difference factors such as socioeconomic status, health, cognition, functional ability, and attitudes, as well as properties of the intervention itself. Adherence research has primarily focused on medication, diet, behavioral therapies, and physical activities (García-Casares et al., 2021; Koutsonida et al., 2021; Kuo et al., 2021, Koesmahargyo et al., 2020; Lo-Ciganic et al., 2015; Kim et al., 2019; Scioscia et al., 2021), and far fewer studies have examined adherence to cognitive training (Sabaté & World Health Organization, 2003; Turunen et al., 2019). Although studies of adherence in other domains provide some basis for the understanding of factors that might predict cognitive training adherence, there may also be unique factors important for determining adherence to this type of intervention specifically. First, cognitive training has different requirements and challenges compared to other health interventions. Moreover, most cognitive training is delivered via computer, tablet, or smartphone technology, which may present additional barriers to long-term engagement, especially for older adults who may be less technology proficient (Roque & Boot, 2018). Therefore, it is unclear whether similar or unique factors might predict older adults’ adherence to technology-based cognitive training (Munro et al., 2007; Picorelli et al., 2014; Rivera-Torres et al., 2019; Room et al., 2017; Sabaté & World Health Organization, 2003; Turunen et al., 2019). A detailed understanding of these factors would make it possible to predict adherence and adherence lapses and consequently maximize intervention effectiveness.

Our on-going “Adherence Promotion with Person-Centered Technology (APPT)” project aims at understanding factors that predict adherence to cognitive training and testing programs, and developing just-in-time AI-based interventions that can predict and prevent adherence failures (Charness et al., 2021). The long-term goal of the project is to use technology-based cognitive assessment and training programs to detect and delay the onset of MCI and Alzheimer’s disease and related dementias. The intervention under development is a reminder system that will send tailored messages to individuals who are at risk for adherence failure at an optimal time for reminders to be effective. A promising approach to developing adherence failure prediction models is to employ state-of-the-art machine learning and deep learning techniques, then use these predictions to deliver just-in-time adherence support, such as motivational messages delivered at a time when they might be most effective.

1.1. Research Objectives

In the previous study (Harrell et al., 2021), participants were asked to engage in the Mind Frontier (Aptima®) cognitive training program in two phases - a 12-week structured phase during which participants were asked to engage with the program according to a pre-defined schedule and a 6-week unstructured phase during which participants could train at their will. The intervention was delivered via 10-inch Lenovo tablets, and participants received training on both the use of the tablet and each game in the laboratory before taking the tablet home to train. A booklet with game and device instructions was also provided to each participant. Researchers examined whether different motivational messages delivered to participants would boost adherence to a technology-based cognitive training program. The message manipulation of interest was largely ineffective, leading the researchers to suggest that individual difference factors may drive adherence behaviors. Previously, this study utilized generalized linear mixed-effects models (GLMM) to predict adherence, defined as the number of game sessions completed.

In the current study, we leveraged data from this previous home-based cognitive intervention and explored the use of machine learning and deep learning approaches for predicting adherence to the cognitive training games. The dataset contains demographic information, measures that assess objective cognition, subjective cognition, functional ability, self-efficacy, technology proficiency, perceived benefits of cognitive training, as well as fine-grained game play data. In this study, newly defined measures of adherence considered not just the number of sessions completed, but also whether those sessions were of appropriate duration as recommended by the experimenters. We defined overall and weekly adherence at minimal and full levels and built machine learning and deep learning models to predict overall adherence of the 12-week structured phase and weekly adherence at both minimal and full levels. We formulated two research questions (RQs):

RQ 1: Can machine learning models with the demographics, attitudinal, and cognitive ability measures collected at the beginning of the trial predict overall adherence to the technology-based cognitive training program? Which variables are more predictive than others?

RQ 2: Can deep learning models with the data collected from daily training interactions in the current week predict the following week's adherence to the technology-based cognitive training program? Which variables are more predictive than others?

The current study's contributions to the body of knowledge are summarized as follows.

1. Machine learning has been previously used to predict adherence to medications (Koesmahargyo et al., 2020; Lo-Ciganic et al., 2015), medical therapies (Kim et al., 2019; Scioscia et al., 2021), and physical activities (Zhou et al., 2019). To the best of our knowledge, this is the first study that uses machine learning to predict adherence to cognitive training.
2. We conducted experiments using different machine learning models with different definitions of adherence to predict adherence to a cognitive training program.

3. We identified salient features that are predictive of both overall adherence and weekly adherence at different levels.
4. The results of this project can be used to facilitate the design of a just-in-time intervention for improving adherence to technology-based cognitive training programs.

The article is organized as follows. In Section 2, we briefly review some related work on adherence to mobile app interventions, adherence to cognitive training, and the Mind Frontier program used in this study. In Section 3, we present the study design, data preparation, modeling, and interpretation of the models. Section 4 presents the experimental results. In Section 5, we discuss the principal findings, their practical implications to adherence promotion, and future directions, followed by conclusions in Section 6.

2. Related Work and Background

2.1. Adherence Promotion and Participant Engagement with Mobile App Interventions

In recent years, healthcare has been undergoing a transformation towards integrating digital technologies including mobile apps, health monitors, and patient portals, resulting in enhanced patient engagement and improved patient outcomes (Dubad et al., 2018). For example, studies have found that patients who use web-based patient portals for diabetes had improve glycemic control (Lau et al., 2014; Osborn et al., 2010). Use of electronic health records has been shown to improve patient activation and empowerment among HIV (Crouch et al., 2015) and cardiac patients (Toscos et al., 2016). However, only a few studies have assessed the effectiveness of different mobile app components on enhancing engagement and, consequently, on improving health outcomes. Oakey-Girvan et al. (2022) conducted a review of studies that have examined app components associated with increased engagement. Measures of engagement can be categorized into quantitative measures which are test-based, and qualitative measures which are based on subjectivity (i.e., participant opinions). Across digital tools for different health conditions, three major types of components were found to be associated with increased engagement and usage: (1) diaries and feedback, (2) coaching and education, (3) reminders (e.g., app notifications, text reminders). In particular, text messages or push notifications have been shown to improve the use of apps for alcohol/substance abuse (Puddephatt et al., 2019; Westergaard et al., 2017), hypertension (Toro-Ramos et al., 2017), mental health (Bidargaddi et al., 2018), and weight loss (Dolan et al., 2019; Morrison et al., 2014). The review did not identify any studies that have explored strategies and best practices for sending the reminders. Our study aims to facilitate the development of effective strategies for sending timely reminders to promote adherence in mobile app interventions.

2.2. Adherence to Cognitive Training

Even though a number of studies have analyzed the effects of cognitive training including the potential to delay the onset of dementia (Bahar-Fuchs et al., 2013; Basak et al., 2020; Chiu et al., 2017; Ge et al., 2018; Gray et al., 2021; Kallio et al., 2017; Nguyen et al., 2019; Sala et al., 2019; Simons et al., 2016; Turunen et al., 2019; Zhang et al., 2019), studies that have evaluated factors correlated with participants' adherence to cognitive

training are scarce. Recently, Harrell et al. (2021) evaluated the effects of message framing on adherence to a home-based cognitive training intervention, and also examined factors that may correlate with participants' willingness to participate in cognitive intervention (e.g., age, belief in the efficacy of cognitive training, previous computer use, self-perceived cognition, technology self-efficacy, and memory performance). Using GLMM, they found that the odds of daily engagement decreased as a function of self-efficacy composite score and found no effects of message framing (i.e., positive-framed vs. negative-framed messages) on participant's adherence to the training in the first 12-week structured phase. Turunen et al. (2019) conducted a study of a computer-based cognitive training (CCT) with participants 60–77-year-old with increased dementia risk. Using zero inflated negative binomial regression analyses, they found that previous experience with computers, being married or co-living, better memory, and positive expectations of the study were correlated with higher chance of starting CCT but previous computer use was the only factor associated with adherence in terms of the number of completed CCT sessions.

3. Materials and Methods

3.1. The Cognitive Training Intervention and the Dataset

The current study analyzes data collected in a previously published study (Harrell et al., 2021). In this earlier study, 118 community-dwelling older adults (*Mean age* = 72.6 years, *SD* = 5.54) were asked to play a series of gamified neuropsychological tasks on tablets in their homes. The study used the Mind Frontiers cognitive training program. Like many commercially available cognitive training packages, the Mind Frontiers software package consists of games modeled after classic measures of memory, attention, spatial processing, task-switching, reasoning ability, and problem-solving within the psychological literature. These gamified neuropsychological tasks were developed to fit within a “Wild West” theme. Feedback is provided after each game, and each game adapts in difficulty according to participants' performance.

The dataset collected in the previous study (Harrell et al., 2021) contained details of 205,002 training interactions (game play sessions) for 118 participants, as well as a host of individual difference measures (see Table 1 for demographics). Multiple objective cognitive measures each assessed the constructs of processing speed, memory, and reasoning. Multiple measures also assessed subjective/self-reported cognitive and functional ability. Attitudinal measures included multiple measures, each tapping constructs of self-efficacy, technology proficiency, and perceived benefits of cognitive training (see Table 2). All measures were z-scored, and composite scores for each construct were also created by averaging measures assessing the same or similar constructs (Table 2). Detailed information about training interactions included training session date, timestamp, duration, task (which training game), task level, and outcome for each participant interaction (Table 3). The cognitive training program featured 7 different tasks/games, and the data include max task levels achieved, ranging from 16 to 58 for different tasks, and 5 outcomes (i.e., defeat, stalemate, victory, abort, not yet finished).

3.2. Overall Design

The current study leverages data collected during a previously published study which is described in more detail in Section 3.1. Figure 1 illustrates the overall design of this study. We first investigated if demographic, attitudinal, and cognitive ability measures are predictive of overall adherence throughout the structured (12 week) phase (similar to Harrell et al., 2021). Then, to predict weekly adherence (using the previous week's game interaction data), we derived a set of variables based on the daily interaction with the games including the time of training, number of sessions, and number of different outcomes, organized them in a longitudinal format in a week, and evaluated multiple deep learning models that can handle time series data. After the best models were found, we applied a post hoc interpretability method to explain the models by ranking the features based on their importance to the predictions.

3.3. Measuring Adherence

The trial consisted of a structured phase of 12 weeks in which participants were asked to follow a predefined training schedule (5 sessions on different days per week, 45 minutes per session) and an unstructured phase of 6 weeks in which participants could engage with training at will. Here we focused only on the structured phase as adherence cannot be defined without a suggested schedule of gameplay. As a first step in calculating adherence, we first assessed if individual training days met the requirements of minimal adherence or full adherence. Training days involving gameplay for at least 10 minutes were considered to satisfy the minimal requirement, and days involving gameplay for at least 36 minutes (80% of suggested time) were considered to satisfy the full requirement. *Weekly adherence at a minimal level* was defined as the number of training days participants satisfied the minimal requirement in a week divided by 5 days; and *weekly adherence at full level* was defined as the number of training days participants satisfied the full requirement in a week divided by 5 days. In other words, these numbers represent the proportion of completed sessions each week relative to the suggested number of sessions recommended by the experimenter, using a loose or strict criteria for what counts as a completed session. Note that weekly adherence at minimal or full level of a participant can be greater than 1, if he/she satisfied the minimal or full requirements for more than 5 days in that week. Averages of weekly adherence at minimal and full levels for each participant were also computed as the *overall adherence* for the structured time of 12 weeks. Participants were then divided into two classes (i.e., low adherence, high adherence) using the median of overall adherence at each threshold.

3.4. Predicting Overall Adherence

To investigate if individual difference measures are predictive of overall adherence, we conducted t-test, and Chi-square test where applicable to examine the associations between overall adherence and demographic, attitudinal, and cognitive measures, and select variables with statistical significance for building machine learning models. Table 4 shows the correlations between the variables and the overall adherence at the minimal and full levels. Continuous variables and categorical variables were selected with p-values at significance level of 0.10 in terms of t-test and Chi-square respectively. No demographic variables were selected by the tests. Seven z-score variables and three composite score variables were

selected as predictors of overall adherence at the minimal level while six z-score variables and three composite score variables were selected as predictors of overall adherence at the full level. Classification models of logistic regression (LR), support vector machine (SVM), classification and regression trees (CART) and random forest (RF) were built for prediction of overall adherence classes based on the selected variables. Due to the small sample size ($N=118$) and the nature of the predictors (i.e., demographics, cognitive and memory function, attitudinal features), deep learning models would not be applicable. We implemented all machine learning models using the scikit-learn Python package with default parameter settings. All machine learning models were tested on z-score predictors or composite score predictors.

3.5. Predicting Weekly Adherence

Since we hypothesized that weekly adherence is associated with game play data, we derived a set of variables from daily interaction details of the training for each day in a week to predict weekly adherences at minimal and full levels in the next week. A total of 16 variables were derived as shown in Table 5. In the experiments of weekly adherence prediction, each participant had 11 weeks' adherence levels (not including Week 1) as the prediction outcomes. The sample size is 1,298 (118 participants \times 11 weeks) for each adherence level (i.e., minimal level and full level). Since we organized our weekly adherence prediction data day-by-day as time-series data, we opted to use deep learning models with various architectures which showed superior performance than traditional machine learning models. In this task, we experimented with deep learning models with (1) two Long Short-Term Memory (LSTM) hidden layers, (2) two Recurrent Neural Network with Gated Recurrent Units (GRU) hidden layers, (3) two 1-dimension convolution (Conv1D) hidden layers (CNN), and (4) all the combinations of two different hidden layers of LSTM, GRU and Conv1D. We implemented all deep learning models using Keras of the TensorFlow Python package. After careful hyperparameter tuning, we set the number of units as 64 for the first hidden layer of LSTM or GRU and 32 for the second hidden layer of LSTM or GRU. Parameters for Conv1D included kernel size of 3 and stride of 1. The number of filters for Conv1D was set to 64 for the first hidden layer and 32 for the second hidden layer. We used Rectified Linear Unit (ReLU) activation function for the Conv1D hidden layers. All models were trained for 10 epochs using Adam optimizer with learning rate of 0.001, loss of categorical cross entropy, and batch size of 32.

3.6. Experimental Design and Evaluation

Accuracy, weighted precision, weighted recall, weighted F1 score, area under receiving operating curve (AUROC) and area under prediction recall curve (AUPRC) were used as evaluation metrics for all the models. For overall adherence classification models, we conducted 5-fold cross validation and computed the average value for each evaluation metric for model comparison. For weekly adherence classification models, we randomly selected 80 participants as training data and held the remaining 38 as the test data. We conducted experiments 5 times with different sets of random samples for training and testing and reported the average performance metrics.

3.7. Model Interpretability

Deep learning models are often considered as black boxes where predictions are difficult and sometimes even impossible to interpret for both users and developers. With the growing application of machine learning and deep learning, it becomes increasingly important to understand why a model makes a certain prediction. Vast numbers of studies in the field known as XAI (eXplainable Artificial Intelligence) (Payrovnaziri et al., 2020) have proposed, developed and tested a wide range of approaches including SHAP (SHapley Additive exPlanations), a state-of-the-art unified framework for XAI, for interpreting machine learning and deep learning models. SHAP² uses an additive feature attribution approach to assign an importance value to each feature for a particular prediction by measuring the change in the expected model prediction compared to the base model conditioning on the feature (Lundberg & Lee, 2017). In this study, we used SHAP LinearExplainer to interpret prediction of the logistic regression (LR) models which achieved best predictions for overall adherences, and SHAP KernelExplainer to interpret deep learning models that achieved best predictions for weekly adherence.

4. Results

4.1. Adherence Distribution

Figure 2 illustrates the distributions of overall adherence on the minimal level and full level. Both distributions are bimodal, indicating that the adherence levels of the participants varied significantly. Figure 3 shows the box plot for the weekly adherence on the minimal level and full level for each week. According to the median values, the weekly adherence (on both minimal level and full level) generally decreased from Week 1 to Week 12.

4.2. Overall adherence prediction

For overall adherence prediction, we created four datasets for two adherence levels and two sets of features. The first set of features included z-score features and the other set included composite scores. The machine learning model using logistic regression outperformed other models with an average AUROC of 0.71. Models using composite scores outperformed the ones using individual z-scores. The detailed performance metrics for each model are shown in Table 6.

4.3. Weekly adherence prediction

In each run, the training data had 880 (80×11) instances and the test data had 418 (38×11) instances. Table 7 shows the average evaluation metrics of the 5 runs for the deep learning classification models using derived variables to predict weekly adherence at minimal and full levels. For minimal adherence prediction, LSTM-GRU and GRU-LSTM models yielded the best AUROC of 0.861. For full adherence prediction, LSTM model yielded the best AUROC of 0.844.

²Available on GitHub, <https://github.com/slundberg/shap>

4.4. Interpretation of Best Performing Models by SHAP Values

Figure 4, 5 show the impact of the selected z-score and composite score predictors in terms of mean absolute SHAP value on overall adherence prediction by LR models. Strength of impact for each predictor was indicated by the length of the corresponding bar. For predictions of overall adherences at both minimal and full levels, General self-efficacy had the highest impact while Rey's AVLT, immediate recall had the lowest impact. When examining impact of the composite score predictors, composite score of objective cognition memory delayed recall had the highest impact on prediction of overall adherence at minimal level while the composite score of self-efficacy had the highest impact on prediction of overall adherence at full level.

Figure 6 shows impact of the derived variables in each day of a week ranked by the mean of the SHAP values on weekly adherence prediction by the deep learning models. The value in each cell is the SHAP value for a particular variable in a particular day. The higher the SHAP value of the variable, the higher the impact of such a variable on the prediction. For predictions of adherences at both minimal and full levels, the predictors derived from daily interaction time (sec: # of seconds each day; ses: # of sessions each day) had higher impact compared to other predictors. When examining impact of the daily predictors by days in a week, predictors in early days tended to have higher impact on the predictions. From Day 1 to Day 7, the impact of sec slightly decreased. The impact of # of sessions stayed the same from Day 1 to Day 5 and decreased in Day 6 and Day 7.

Figure 7 shows the adherence patterns, breakdown between low adherence and high adherence, entropy, and best model's prediction accuracy of weekly adherence, at both the minimal level and full level for each participant. The entropy was calculated as $-\sum_i P_i \log(P_i)$, where P_i is the proportion of low adherence and high adherence. It can be used to measure the consistency of adherence pattern. As shown in the figure, it is obvious that participants with a consistent adherence pattern (mostly low adherence or mostly high adherence, e.g., 60ADS, 53QQQ) had lower entropy and high prediction accuracy. In contrast, those with less consistent adherence pattern (e.g., 60XXX, 11AAA) had higher entropy and lower prediction accuracy.

5. Discussion

5.1. Main Findings

Although the effect of cognitive training on cognitive functions and dementia among older adults has been mixed in the literature (Bahar-Fuchs et al., 2013; Basak et al., 2020; Chiu et al., 2017; Ge et al., 2018; Kallio et al., 2017; Nguyen et al., 2019; Sala et al., 2019; Simons et al., 2016; Turunen et al., 2019; Zhang et al., 2019), findings from recent prominent studies such as the Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE) study (Rebok et al., 2014) have shown that cognitive training may hold promise for improving cognition and preserving functional abilities. Other studies (Jaeggi et al., 2011) have suggested that limiting factors such as individual differences in training performance and training regimen should be considered when evaluating the effects of the training. However, a limiting factor in determining the efficacy of cognitive training can be

poor intervention adherence (Boot et al., 2013; Corbett et al., 2015; Turunen et al., 2019). Adherence is a key issue in determining cognitive training efficacy, and in ensuring benefits are realized in the population should definitive efficacy be demonstrated.

To the best of our knowledge, this is the first study that uses machine learning to predict adherence to cognitive training programs. We made several original contributions to the understanding of adherence to cognitive training and experimental design of cognitive training programs. First, in this project, we defined different levels of adherence (i.e., minimal level and full level) and analyzed the association between demographics, attitudinal, and cognitive ability measures of participants in a cognitive training trial and their overall adherence to the training program at different levels. The results provide insights into the factors that may influence the overall adherence at different levels with implications on how to develop customized strategies for adherence promotion and on how to stratify participants for future clinical trials designed to test the effectiveness of cognitive training.

Second, we also used machine learning to predict weekly adherence to the training program. Specifically, we were able to use prediction results to determine when a participant's adherence would drop over the course of the intervention. This data can be used to inform the development and delivery of just-in-time adherence promotion strategies to support cognitive training interventions.

According to the findings, general self-efficacy, objective cognitive and memory measures with delayed recall, and technology self-efficacy are most predictive of a participant's overall adherence. Various studies have also shown a strong correlation between self-efficacy, treatment adherence and intervention outcomes (Kelly et al., 2014; Okuboyejo et al., 2018; Turunen et al., 2019). Although these were also observed in this dataset using GLMM to predict adherence in terms of number of days on which participants played a game session by Harrell et al. (2021), we also observed that technology readiness and belief in cognitive training (e.g., NICT) may also be important factors to consider. Regarding weekly adherence prediction, the models predicting adherence at both minimal level and full level achieved high performance. Although intuitively, we would have predicted that participants experiencing more wins than losses might persist more – and show higher adherence – the best predictor of engagement in the intervention next week was simply engagement this week – the amount of time of gameplay. We also found that participants with high entropy and inconsistent adherence patterns have lower prediction accuracy of weekly adherence than those with lower entropy and consistent adherence patterns. This subgroup warrants further investigations for adherence prediction and promotion.

5.2. Theoretical and Practical Implications for Adherence Promotion

To be effective, computer and mobile-based cognitive training interventions require consistent user involvement over time. This may be particularly challenging for older adults to sustain for a variety of reasons such as low computer self-efficacy (Czaja et al., 2006). Adherence to technology-based health interventions has been a persistent problem in a variety of health contexts with high attrition and high disengagement rates commonly being reported in various studies (Wei et al., 2020). In their systematic

review, Wei et al. (2020) identified several design features that influence engagement in mHealth interventions: personalization, reinforcement, communication, navigation, credibility, message presentation, and interface aesthetics, which in turn, can influence adherence to directives of the health intervention aimed at obtaining desired distal health outcomes. In our study, we examined factors that contribute to older adults' consistent engagement with mobile-based cognitive training games, which can then help inform the development and implementation of design features to improve overall adherence to cognitive training.

Moreover, a growing number of mHealth intervention studies have been exploring how app use patterns may influence engagement. Alshurafa et al. (2018) studied the effects of an mHealth intervention designed to improve selected health behaviors among young adults (N=303). In the study, participants' engagement and app use patterns were tracked over a one-year period. They then analyzed the relationship between different incentive strategies deployed at different times, and engagement and app use patterns, and how these, in turn, correlated with health behaviors. The researchers were able to identify different patterns of interaction with the intervention and suggested different strategies for incentivizing continued engagement. They also proposed that effective engagement over time can be improved by considering use patterns.

Our study builds on this strategy by using machine learning to analyze use patterns, identify periods of high or low engagement, predict future use patterns on the fly, and use this data to deliver appropriate engagement strategies. In particular, we intend to use machine learning to predict when to send "smart" reminders at the most opportune times in order to promote adherence to the training programs. Learning and cognitive theories emphasize the role of shaping a desired behavior by identifying natural opportunities for learning and improvement (Lawson & Flocke, 2009) and immediately reinforcing the target behavior when acquiring a new skill (Krueger & Dayan, 2009). These theories indicate that providing timely intervention scaffolds and prompts can capitalize on short-term natural opportunities to improve health outcomes.

In terms of theoretical advancement of models of adherence and health behaviors, we believe that the approaches outlined in this paper have the potential to make an important contribution. As summarized by Rejeski and Fanning (2019), traditional theories often fail to consider "in-the-moment determinants of health behavior." Adherence support strategies that anticipate current, in-the-moment, receptiveness to intervention engagement based on recent history are more likely to be successful, compared to strategies guided by static or staged health behavior models. The machine learning approaches outlined here recognize that adherence is a behavior that continuously unfolds over time and that theories that account for moment-to-moment determinants of health behaviors may be more fruitful in understanding fundamental determinants of adherence and adherence lapses. Machine learning approaches applied to longitudinal data are likely to play an important role in developing such theories.

In our on-going "Adherence Promotion with Person-Centered Technology" (APPT) project, we aim to address the problem of low adherence to mobile-based cognitive training by

developing an AI reminder system to support homebased assessments and interventions, with the ultimate goal of promoting early detection and treatment of age-related declines in cognition. Using results of our research on AI-based adherence prediction, we will design a Just-in-time adaptive intervention (JITAI) in the form of a tailored reminder system to encourage our study participants to follow the training schedule with the Mind Frontiers mobile-based cognitive training game suite. JITAI refers to intervention designs that adapt strategies and the delivery of supports based on an individual's changing status and contexts over time (Nahum-Shani et al., 2018). Specifically, instead of sending generic, one-size-fits-all messages for all participants at a fixed time, our system will be designed to tailor messages based on individual differences, such as personality and motivation to participate in the study. It will also adjust the timing of the messages based on when each participant usually engages with the training and assessments and when the participant's adherence would likely drop. In other parallel studies, we are investigating older adult's motivation for participating in cognitive training (Carr et al., 2022) and the effectiveness of tailored messages for promoting cognitive training research. These studies will provide fruitful insights into the design of AI-based adherence promotion systems for older adults.

5.3. Limitation and Future Work

Several limitations should be noted in this study. First, the sample size is relatively small and there were significantly more females than males in the study (i.e., 78 females vs. 38 males), which may impact generalizability of the reported results. The models we developed in this study are specific to our dataset. We will further conduct work to demonstrate the generalizability of the models to other domains. Nonetheless, we believe that the framework of using machine learning to predict adherence at a certain interval can be generalized to other mobile interventions. Second, certain gameplay data such as play time may not be accurate as participants may have put the tablet aside while the game was still running. In future work, we will optimize the machine learning models for those participants with high entropy. We will cluster the participants based on their weekly adherence data and identify the subgroups with similar adherence patterns. We will then analyze the correlations of the demographic, attitudinal, and cognitive abilities to the adherence patterns and develop more specialized machine learning-based prediction models for specific subgroups of participants with a certain adherence pattern.

6. Conclusions

In this work, we used data from a previous trial of cognitive training for older adults to evaluate the feasibility of using machine learning and deep learning algorithms to predict cognitive training adherence. We defined weekly and overall adherence at two levels. Machine learning classification models were developed to predict overall adherence based on attitudinal and cognitive ability measures selected through statistical analyses. We also derived variables from daily training interactions and built deep learning models for predicting the next week's adherence. Our experiments showed that machine learning models based on selected variables of attitudinal and cognitive ability measures can only achieve moderate accuracy for predicting overall adherence, while the deep learning models

based on variables derived from daily interactions were able to predict weekly adherence with high accuracy.

Acknowledgments

This work was supported by the National Institute on Aging grants R01AG064529 and 4P01AG17211, as well as University of Florida-Florida State University Clinical and Translational Science Award UL1TR001427 supported by National Center for Advancing Translational Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

8. References

- Administration on Aging. (2020). 2020 Profile of Older Americans. Retrieved from https://acl.gov/sites/default/files/Aging%20and%20Disability%20in%20America/2020ProfileOlderAmericans.Final_.pdf
- Allaire JC, & Marsiske M (1999). Everyday Cognition: Age and Intellectual Ability Correlates. *Psychology and Aging*, 14(4), 627–644. [PubMed: 10632150]
- Alshurafa N, Jain J, Alharbi R, Iakovlev G, Spring B, & Pfammatter A (2018). Is More Always Better? Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies. 10.1145/3287031
- Arthur W Jr, & Day DV (1994). Development of a short form for the Raven Advanced Progressive Matrices Test. *Educational and Psychological Measurement*, 54(2), 394–403.
- Bahar-Fuchs A, Clare L, & Woods B (2013). Cognitive training and cognitive rehabilitation for persons with mild to moderate dementia of the Alzheimer's or vascular type: A review. *Alzheimer's Research & Therapy*, 5(4), 35. 10.1186/alzrt189
- Basak C, Qin S, & O'Connell MA (2020). Differential effects of cognitive training modules in healthy aging and mild cognitive impairment: A comprehensive meta-analysis of randomized controlled trials. *Psychology and Aging*, 35(2), 220–249. 10.1037/pag0000442 [PubMed: 32011155]
- Berry JM, West RL, & Dennehey DM (1989). Reliability and validity of the Memory Self-Efficacy Questionnaire. *Developmental Psychology*, 25(5), 701.
- Bidargaddi N, Almirall D, Murphy S, Nahum-Shani I, Kovalcik M, Pituch T, Maaieh H, & Strecher V (2018). To Prompt or Not to Prompt? A Microrandomized Trial of Time-Varying Push Notifications to Increase Proximal Engagement With a Mobile Health App. *JMIR MHealth and UHealth*, 6(11), e10123. 10.2196/10123 [PubMed: 30497999]
- Boot W, Champion M, Blakely D, Wright T, Souders D, & Charness N (2013). Video Games as a Means to Reduce Age-Related Cognitive Decline: Attitudes, Compliance, and Effectiveness. *Frontiers in Psychology*, 4. 10.3389/fpsyg.2013.00031
- Boot WR, Charness N, Czaja SJ, Sharit J, Rogers WA, Fisk AD, Mitzner T, Lee CC, & Nair S (2015). Computer Proficiency Questionnaire: Assessing Low and High Computer Proficient Seniors. *The Gerontologist*, 55(3), 404–411. 10.1093/geront/gnt117 [PubMed: 24107443]
- Brandt J, & Benedict RH (2001). Hopkins verbal learning test–revised: Professional manual. Psychological Assessment Resources.
- Carr DC, Tian S, He Z, Chakraborty S, Dieciuc M, Gray N, Agharazidermani M, Lustria MLA, Dilanchian A, Zhang S, Charness N, Terracciano A, & Boot WR (2022). Motivation to Engage in Aging Research: Are There Typologies and Predictors? *The Gerontologist*, gnac035. 10.1093/geront/gnac035 [PubMed: 35267020]
- Charness N, Boot W, Carr D, Chakraborty S, He Z, Lustria M, & Terracciano A (2021). Aims of the Adherence Promotion With Person-Centered Technology (APPT) Project. *Innovation in Aging*, 5(Supplement_1), 551. 10.1093/geroni/igab046.2116
- Chiu H-L, Chu H, Tsai J-C, Liu D, Chen Y-R, Yang H-L, & Chou K-R (2017). The effect of cognitive-based training for the healthy older people: A meta-analysis of randomized controlled trials. *PLOS ONE*, 12(5), e0176742. 10.1371/journal.pone.0176742 [PubMed: 28459873]
- Corbett A, Owen A, Hampshire A, Grahn J, Stenton R, Dajani S, Burns A, Howard R, Williams N, Williams G, & Ballard C (2015). The Effect of an Online Cognitive Training Package in Healthy

- Older Adults: An Online Randomized Controlled Trial. *Journal of the American Medical Directors Association*, 16(11), 990–997. 10.1016/j.jamda.2015.06.014 [PubMed: 26543007]
- Crouch P-CB, Rose CD, Johnson M, & Janson SL (2015). A pilot study to evaluate the magnitude of association of the use of electronic personal health records with patient activation and empowerment in HIV-infected veterans. *PeerJ*, 3, e852. 10.7717/peerj.852 [PubMed: 25802815]
- Czaja SJ, Charness N, Fisk AD, Hertzog C, Nair SN, Rogers WA, & Sharit J (2006). Factors Predicting the Use of Technology: Findings From the Center for Research and Education on Aging and Technology Enhancement (CREATE). *Psychology and Aging*, 21(2), 333–352. 10.1037/0882-7974.21.2.333 [PubMed: 16768579]
- Diehl M, Willis SL, & Schaie KW (1995). Everyday problem solving in older adults: Observational assessment and cognitive correlates. *Psychology and Aging*, 10(3), 478–491. 10.1037//0882-7974.10.3.478 [PubMed: 8527068]
- Dolan PT, Afaneh C, Dakin G, Pomp A, & Yeo HL (2019). Lessons Learned From Developing a Mobile App to Assist in Patient Recovery After Weight Loss Surgery. *Journal of Surgical Research*, 244, 402–408. 10.1016/j.jss.2019.06.063 [PubMed: 31325662]
- Dubad M, Winsper C, Meyer C, Livanou M, & Marwaha S (2018). A systematic review of the psychometric properties, usability and clinical impacts of mobile mood-monitoring applications in young people. *Psychological Medicine*, 48(2), 208–228. 10.1017/S0033291717001659 [PubMed: 28641609]
- Edwards JD, Ross LA, Wadley VG, Clay OJ, Crowe M, Roenker DL, & Ball KK (2006). The useful field of view test: Normative data for older adults. *Archives of Clinical Neuropsychology*, 21(4), 275–286. [PubMed: 16704918]
- Ekstrom RB, & Harman HH (1976). Manual for kit of factor-referenced cognitive tests, 1976. Educational testing service.
- Foldes SS, Moriarty JP, Farseth PH, Mittelman MS, & Long KH (2018). Medicaid Savings From The New York University Caregiver Intervention for Families with Dementia. *The Gerontologist*, 58(2), e97–e106. 10.1093/geront/gnx077 [PubMed: 29106531]
- García-Casares N, Gallego Fuentes P, Barbancho MÁ, López-Gigosos R, García-Rodríguez A, & Gutiérrez-Bedmar M (2021). Alzheimer’s Disease, Mild Cognitive Impairment and Mediterranean Diet. A Systematic Review and Dose-Response Meta-Analysis. *Journal of Clinical Medicine*, 10(20), 4642. 10.3390/jcm10204642 [PubMed: 34682764]
- Ge S, Zhu Z, Wu B, & McConnell ES (2018). Technology-based cognitive training and rehabilitation interventions for individuals with mild cognitive impairment: A systematic review. *BMC Geriatrics*, 18. 10.1186/s12877-018-0893-1
- Gray N, Yoon J-S, Charness N, Boot WR, Roque NA, Andringa R, Harrell ER, Lewis KG, & Vitale T (2021). Relative effectiveness of general versus specific cognitive training for aging adults. *Psychology and Aging*. 10.1037/pag0000663
- Harrell ER, Roque NA, Boot WR, & Charness N (2021). Investigating Message Framing to Improve Adherence to Technology-Based Cognitive Interventions. *Psychology and Aging*, 36(8), 974–982. 10.1037/pag0000629 [PubMed: 34460281]
- Harrell ER, Kmetz B, & Boot WR (2019). Is Cognitive Training Worth It? Exploring Individuals’ Willingness to Engage in Cognitive Training. *Journal of Cognitive Enhancement : Towards the Integration of Theory and Practice*, 3(4), 405–415. 10.1007/s41465-019-00129-4 [PubMed: 31773088]
- Jaeggi SM, Buschkuhl M, Jonides J, & Shah P (2011). Short- and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences*, 108(25), 10081–10086. 10.1073/pnas.1103228108
- Kallio E-L, Öhman H, Kautiainen H, Hietanen M, & Pitkälä K (2017). Cognitive Training Interventions for Patients with Alzheimer’s Disease: A Systematic Review. *Journal of Alzheimer’s Disease*, 56(4), 1349–1372. 10.3233/JAD-160810
- Kelly ME, Loughrey D, Lawlor BA, Robertson IH, Walsh C, & Brennan S (2014). The impact of cognitive training and mental stimulation on cognitive and everyday functioning of healthy older adults: A systematic review and meta-analysis. *Ageing Research Reviews*, 15, 28–43. 10.1016/j.arr.2014.02.004 [PubMed: 24607830]

- Kim N, McCarthy DE, Loh W-Y, Cook JW, Piper ME, Schlam TR, & Baker TB (2019). Predictors of adherence to nicotine replacement therapy: Machine learning evidence that perceived need predicts medication use. *Drug and Alcohol Dependence*, 205, 107668. 10.1016/j.drugalcdep.2019.107668 [PubMed: 31707266]
- Koesmahargyo V, Abbas A, Zhang L, Guan L, Feng S, Yadav V, & Galatzer-Levy IR (2020). Accuracy of machine learning-based prediction of medication adherence in clinical research. *Psychiatry Research*, 294, 113558. 10.1016/j.psychres.2020.113558 [PubMed: 33242836]
- Koutsonida M, Kanellopoulou A, Markozannes G, Gousia S, Doulas MT, Sigounas DE, Tzouvaras VT, Vakalis K, Tzoulaki I, Evangelou E, Rizos EC, Ntzani E, Aretouli E, & Tsilidis KK (2021). Adherence to Mediterranean Diet and Cognitive Abilities in the Greek Cohort of Epirus Health Study. *Nutrients*, 13(10), 3363. 10.3390/nu13103363 [PubMed: 34684367]
- Kramer AF, & Colcombe S (2018). Fitness Effects on the Cognitive Function of Older Adults: A Meta-Analytic Study—Revisited. *Perspectives on Psychological Science*, 13(2), 213–217. 10.1177/1745691617707316 [PubMed: 29592650]
- Krueger KA, & Dayan P (2009). Flexible shaping: How learning in small steps helps. *Cognition*, 110(3), 380–394. 10.1016/j.cognition.2008.11.014 [PubMed: 19121518]
- Kuo W-Y, Chen M-C, Lin Y-C, Yan S-F, & Shyu Y-IL (2021). Trajectory of adherence to home rehabilitation among older adults with hip fractures and cognitive impairment. *Geriatric Nursing*, 42(6), 1569–1576. 10.1016/j.gerinurse.2021.10.019 [PubMed: 34763234]
- Langa KM, Larson EB, Crimmins EM, Faul JD, Levine DA, Kabeto MU, & Weir DR (2017). A Comparison of the Prevalence of Dementia in the United States in 2000 and 2012. *JAMA Internal Medicine*, 177(1), 51–58. 10.1001/jamainternmed.2016.6807 [PubMed: 27893041]
- Lau M, Campbell H, Tang T, Thompson DJS, & Elliott T (2014). Impact of Patient Use of an Online Patient Portal on Diabetes Outcomes. *Canadian Journal of Diabetes*, 38(1), 17–21. 10.1016/j.cjcd.2013.10.005 [PubMed: 24485208]
- Lawson PJ, & Flocke SA (2009). Teachable moments for health behavior change: A concept analysis. *Patient Education and Counseling*, 76(1), 25–30. 10.1016/j.pec.2008.11.002 [PubMed: 19110395]
- Lawton MP, & Brody EM (1969). Assessment of older people: Self-maintaining and instrumental activities of daily living. *The Gerontologist*, 9(3 part 1), 179–186. [PubMed: 5349366]
- Lo-Ciganic W-H, Donohue JM, Thorpe JM, Perera S, Thorpe CT, Marcum ZA, & Gellad WF (2015). Using Machine Learning to Examine Medication Adherence Thresholds and Risk of Hospitalization. *Medical Care*, 53(8), 720–728. 10.1097/MLR.0000000000000394 [PubMed: 26147866]
- Lundberg SM, & Lee S-I (2017). A unified approach to interpreting model predictions. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 4768–4777.
- Morrison LG, Hargood C, Lin SX, Dennison L, Joseph J, Hughes S, Michaelides DT, Johnston D, Johnston M, Michie S, Little P, Smith PW, Weal MJ, & Yardley L (2014). Understanding Usage of a Hybrid Website and Smartphone App for Weight Management: A Mixed-Methods Study. *Journal of Medical Internet Research*, 16(10), e201. 10.2196/jmir.3579 [PubMed: 25355131]
- Munro S, Lewin S, Swart T, & Volmink J (2007). A review of health behaviour theories: How useful are these for developing interventions to promote long-term medication adherence for TB and HIV/AIDS? *BMC Public Health*, 7(1), 104. 10.1186/1471-2458-7-104 [PubMed: 17561997]
- Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, & Murphy SA (2018). Just-in-Time Adaptive Interventions (JITAs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support. *Annals of Behavioral Medicine*, 52(6), 446–462. 10.1007/s12160-016-9830-8 [PubMed: 27663578]
- Nguyen L, Murphy K, & Andrews G (2019). Immediate and long-term efficacy of executive functions cognitive training in older adults: A systematic review and meta-analysis. *Psychological Bulletin*, 145(7), 698–733. 10.1037/bul0000196 [PubMed: 30998045]
- Oakley-Girvan I, Yunis R, Longmire M, & Ouillon JS (2022). What Works Best to Engage Participants in Mobile App Interventions and e-Health: A Scoping Review. *Telemedicine and E-Health*, 28(6), 768–780. 10.1089/tmj.2021.0176 [PubMed: 34637651]

- Okuboyejo S, Mbarika V, & Omoregbe N (2018). The effect of self-efficacy and outcome expectation on medication adherence behavior. *Journal of Public Health in Africa*, 9(3), 826. 10.4081/jphia.2018.826 [PubMed: 30854176]
- Osborn CY, Mayberry LS, Mulvaney SA, & Hess R (2010). Patient Web Portals to Improve Diabetes Outcomes: A Systematic Review. *Current Diabetes Reports*, 10(6), 422–435. 10.1007/s11892-010-0151-1 [PubMed: 20890688]
- Parasuraman A, & Colby CL (2015). An updated and streamlined technology readiness index: TRI 2.0. *Journal of Service Research*, 18(1), 59–74.
- Payrovnaziri SN, Chen Z, Rengifo-Moreno P, Miller T, Bian J, Chen JH, Liu X, & He Z (2020). Explainable artificial intelligence models using real-world electronic health record data: A systematic scoping review. *Journal of the American Medical Informatics Association : JAMIA*, 27(7), 1173–1185. 10.1093/jamia/ocaa053 [PubMed: 32417928]
- Picorelli AMA, Pereira LSM, Pereira DS, Felício D, & Sherrington C (2014). Adherence to exercise programs for older people is influenced by program characteristics and personal factors: A systematic review. *Journal of Physiotherapy*, 60(3), 151–156. 10.1016/j.jphys.2014.06.012 [PubMed: 25092418]
- Puddephatt J-A, Leightley D, Palmer L, Jones N, Mahmoodi T, Drummond C, Rona RJ, Fear NT, Field M, & Goodwin L (2019). A Qualitative Evaluation of the Acceptability of a Tailored Smartphone Alcohol Intervention for a Military Population: Information About Drinking for Ex-Serving Personnel (InDEX) App. *JMIR MHealth and UHealth*, 7(5), e12267. 10.2196/12267 [PubMed: 31127726]
- Rabipour S, & Davidson PS (2015). Do you believe in brain training? A questionnaire about expectations of computerised cognitive training. *Behavioural Brain Research*, 295, 64–70. [PubMed: 25591472]
- Rafii MS, & Aisen PS (2015). Advances in Alzheimer's Disease Drug Development. *BMC Medicine*, 13(1), 62. 10.1186/s12916-015-0297-4 [PubMed: 25857341]
- Raven JC, & Court JH (1994). *Advanced Progressive Matrices Raven Manual*. Oxford Psychologists Press Oxford.
- Rebok GW, Ball K, Guey LT, Jones RN, Kim H-Y, King JW, Marsiske M, Morris JN, Tennstedt SL, Unverzagt FW, & Willis SL (2014). Ten-Year Effects of the Advanced Cognitive Training for Independent and Vital Elderly Cognitive Training Trial on Cognition and Everyday Functioning in Older Adults. *Journal of the American Geriatrics Society*, 62(1), 16–24. 10.1111/jgs.12607 [PubMed: 24417410]
- Rejeski WJ, & Fanning J (2019). Models and theories of health behavior and clinical interventions in aging: A contemporary, integrative approach. *Clinical Interventions in Aging*, 14, 1007–1019. 10.2147/CIA.S206974 [PubMed: 31213787]
- Rivera-Torres S, Fahey TD, & Rivera MA (2019). Adherence to Exercise Programs in Older Adults: Informative Report. *Gerontology and Geriatric Medicine*, 5. 10.1177/2333721418823604
- Roberts R, & Knopman DS (2013). Classification and Epidemiology of MCI. *Clinics in Geriatric Medicine*, 29(4), 10.1016/j.cger.2013.07.003. 10.1016/j.cger.2013.07.003
- Room J, Hannink E, Dawes H, & Barker K (2017). What interventions are used to improve exercise adherence in older people and what behavioural techniques are they based on? A systematic review. *BMJ Open*, 7(12), e019221. 10.1136/bmjopen-2017-019221
- Roque NA, & Boot WR (2018). A new tool for assessing mobile device proficiency in older adults: The mobile device proficiency questionnaire. *Journal of Applied Gerontology*, 37(2), 131–156. [PubMed: 27255686]
- Sabaté E, & World Health Organization (Eds.). (2003). *Adherence to long-term therapies: Evidence for action*. World Health Organization.
- Sala G, Aksayli ND, Tatlidil KS, Tatsumi T, Gondo Y, & Gobet F (2019). Near and Far Transfer in Cognitive Training: A Second-Order Meta-Analysis. *Collabra: Psychology*, 5(1), 18. 10.1525/collabra.203
- Salthouse TA (2010). *Major issues in cognitive aging* (pp. ix, 246). Oxford University Press.
- Schmidt M (1996). *Rey auditory verbal learning test: A handbook*. Western Psychological Services Los Angeles, CA.

- Schwarzer R, & Jerusalem M (1995). Generalized self-efficacy scale. *Measures in Health Psychology: A User's Portfolio. Causal and Control Beliefs*, 1(1), 35–37.
- Scioscia G, Tondo P, Foschino Barbaro MP, Sabato R, Gallo C, Maci F, & Lacedonia D (2021). Machine learning-based prediction of adherence to continuous positive airway pressure (CPAP) in obstructive sleep apnea (OSA). *Informatics for Health and Social Care*, 0(0), 1–9. 10.1080/17538157.2021.1990300
- Simons DJ, Boot WR, Charness N, Gathercole SE, Chabris CF, Hambrick DZ, & Stine-Morrow EAL (2016). Do “Brain-Training” Programs Work? *Psychological Science in the Public Interest*, 17(3), 103–186. 10.1177/1529100616661983 [PubMed: 27697851]
- Sr PAA, DeFeis B, De Wit L, O’Shea D, Mejia A, Chandler M, Locke DEC, Fields J, Phatak V, Dean PM, Crook J, & Smith G (2020). Functional ability is associated with higher adherence to behavioral interventions in mild cognitive impairment. *The Clinical Neuropsychologist*, 34(5), 937–955. 10.1080/13854046.2019.1672792 [PubMed: 31608773]
- Sullivan MJ, Edgley K, & Dehoux E (1990). A survey of multiple sclerosis: I. Perceived cognitive problems and compensatory strategy use. *Canadian Journal of Rehabilitation*, 4(2), 99–105.
- Toro-Ramos T, Kim Y, Wood M, Rajda J, Niejadlik K, Honcz J, Marrero D, Fawer A, & Michaelides A (2017). Efficacy of a mobile hypertension prevention delivery platform with human coaching. *Journal of Human Hypertension*, 31(12), 795–800. 10.1038/jhh.2017.69 [PubMed: 28972573]
- Toscos T, Daley C, Heral L, Doshi R, Chen Y-C, Eckert GJ, Plant RL, & Mirro MJ (2016). Impact of electronic personal health record use on engagement and intermediate health outcomes among cardiac patients: A quasi-experimental study. *Journal of the American Medical Informatics Association : JAMIA*, 23(1), 119–128. 10.1093/jamia/ocv164 [PubMed: 26912538]
- Turunen M, Hokkanen L, Bäckman L, Stigsdotter-Neely A, Hänninen T, Paajanen T, Soininen H, Kivipelto M, & Ngandu T (2019). Computer-based cognitive training for older adults: Determinants of adherence. *PLOS ONE*, 14(7), e0219541. 10.1371/journal.pone.0219541 [PubMed: 31291337]
- Wechsler D (1997). *Wechsler, Adult Intelligence Scale III: 3rd ed.* The Psychological Corporation, San Antonio, TX.
- Wei Y, Zheng P, Deng H, Wang X, Li X, & Fu H (2020). Design Features for Improving Mobile Health Intervention User Engagement: Systematic Review and Thematic Analysis. *Journal of Medical Internet Research*, 22(12), e21687. 10.2196/21687 [PubMed: 33295292]
- Westergaard RP, Genz A, Panico K, Surkan PJ, Keruly J, Hutton HE, Chang LW, & Kirk GD (2017). Acceptability of a mobile health intervention to enhance HIV care coordination for patients with substance use disorders. *Addiction Science & Clinical Practice*, 12, 11. 10.1186/s13722-017-0076-y [PubMed: 28441962]
- Zhang H, Huntley J, Bhome R, Holmes B, Cahill J, Gould RL, Wang H, Yu X, & Howard R (2019). Effect of computerised cognitive training on cognitive outcomes in mild cognitive impairment: A systematic review and meta-analysis. *BMJ Open*, 9(8). 10.1136/bmjopen-2018-027062
- Zhou M, Fukuoka Y, Goldberg K, Vittinghoff E, & Aswani A (2019). Applying machine learning to predict future adherence to physical activity programs. *BMC Medical Informatics and Decision Making*, 19, 169. 10.1186/s12911-019-0890-0 [PubMed: 31438926]

Overall Adherence Prediction

Weekly Adherence Prediction

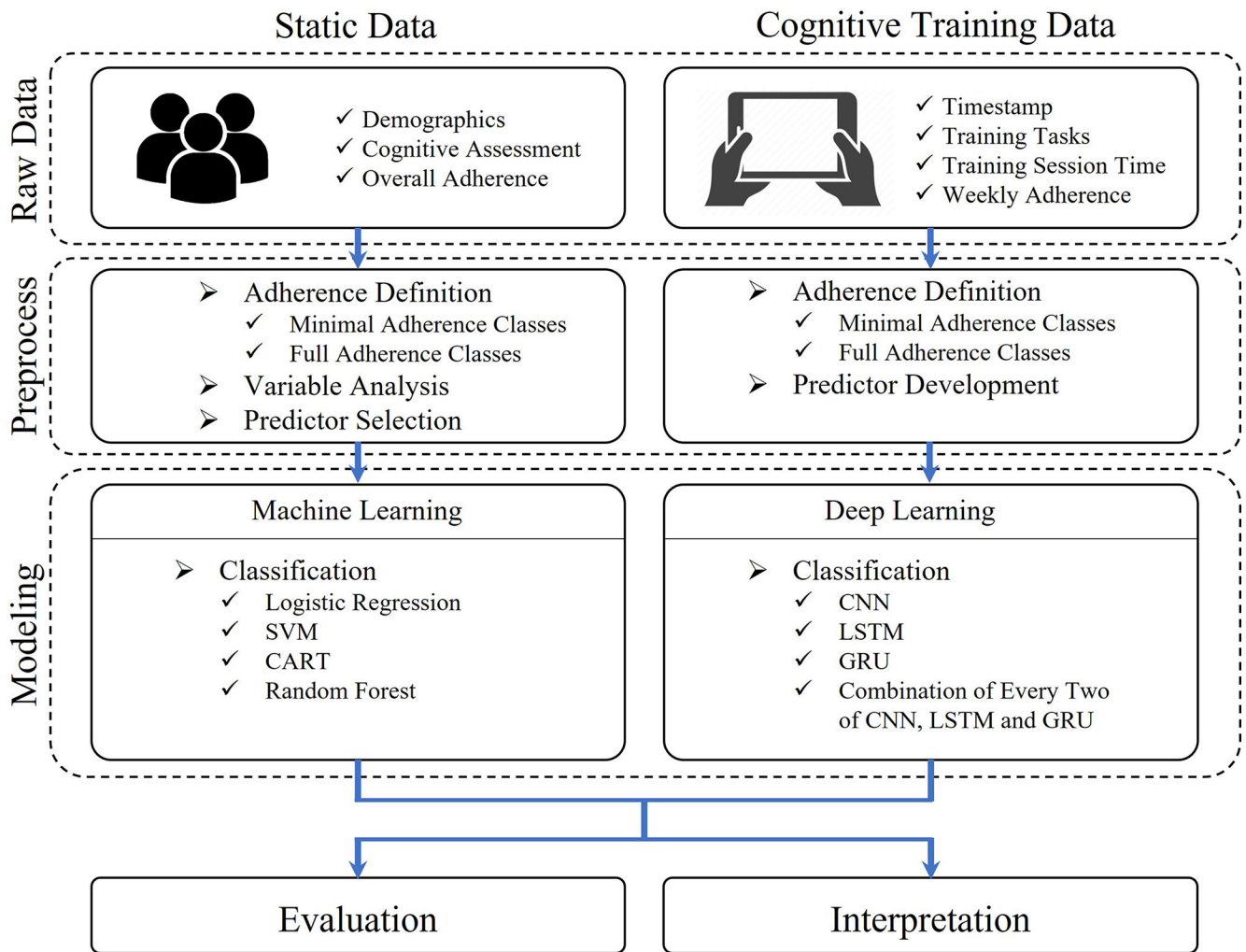


Figure 1. Illustration of the overall design of the study.

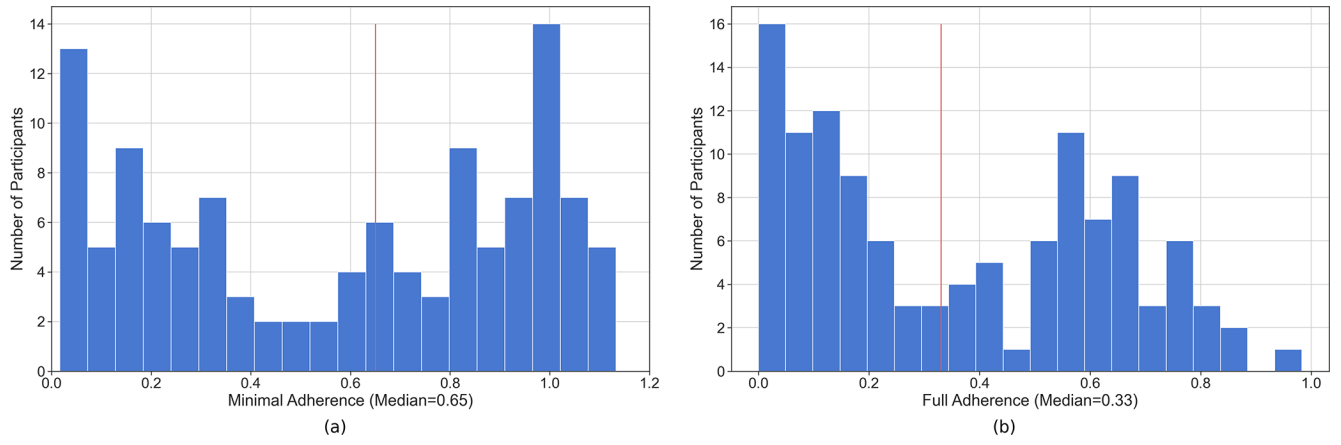


Figure 2. Distribution of overall adherence at (a) minimal level; and (b) full level.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

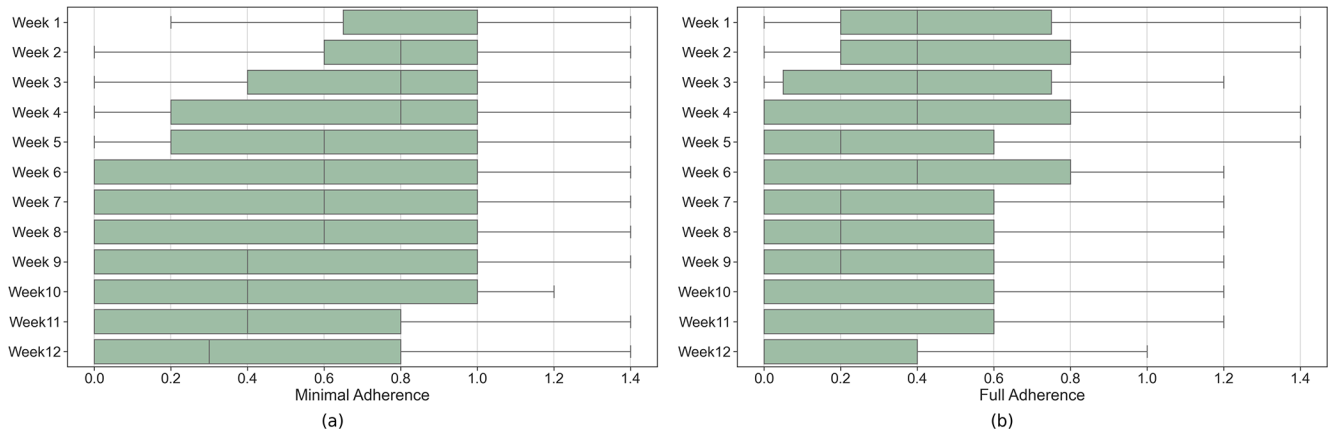


Figure 3. Distribution of weekly adherence at the (a) minimal level and (b) full level.

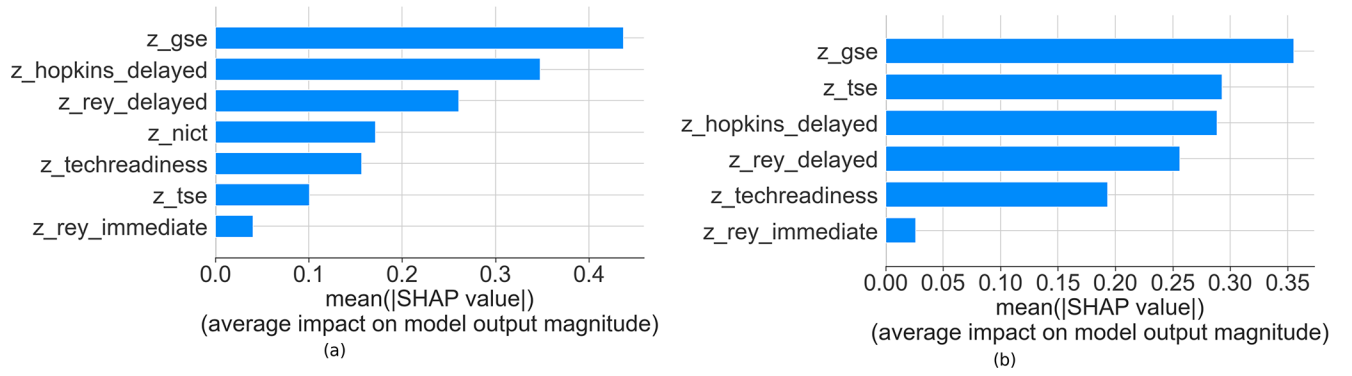


Figure 4. Impact of the selected z-score predictors by mean SHAP value for LR for predicting (a) overall adherence at minimal level and (b) overall adherence at full level.

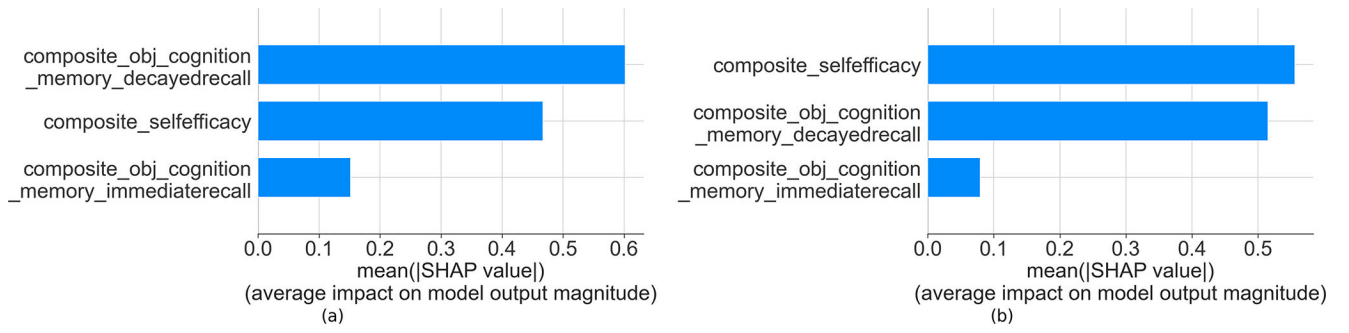


Figure 5. Impact of the selected composite score predictors by mean SHAP value for LR for predicting (a) overall adherence at minimal level and (b) overall adherence at full level.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

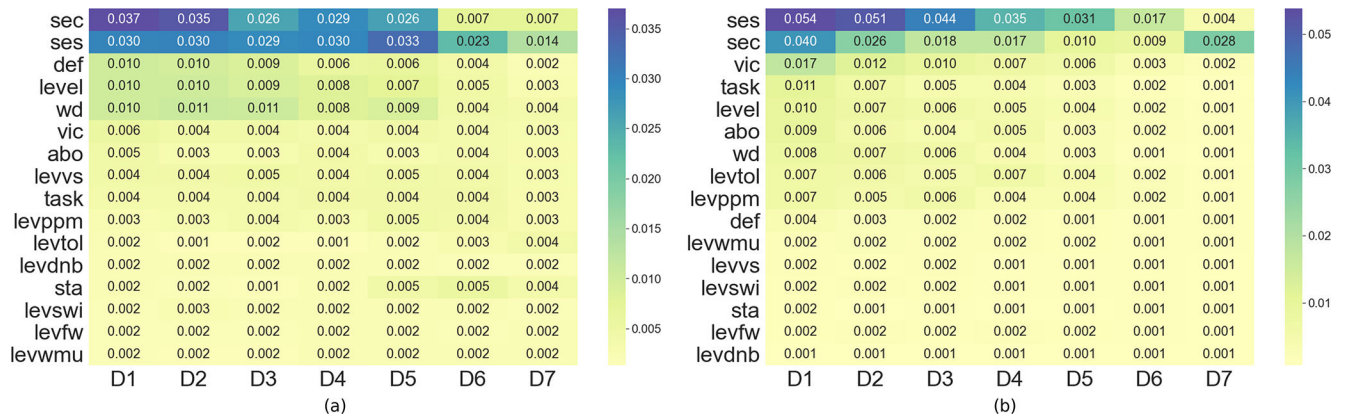


Figure 6. Impact of the derived variables for each day in a week by mean SHAP value for predicting (1) weekly adherence at minimal level with LSTM-GRU and (b) predicting weekly adherence at full level by LSTM.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

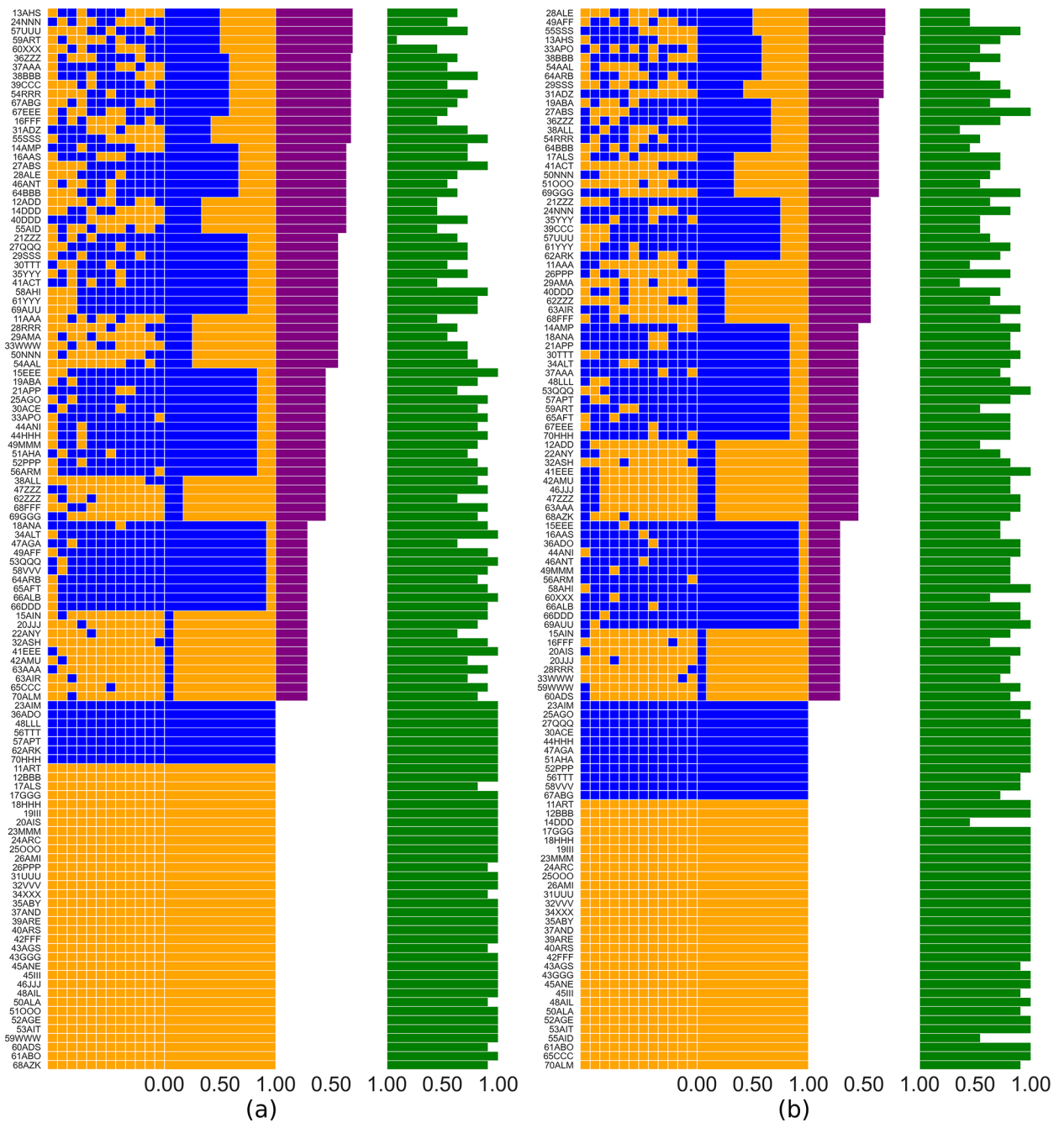


Figure 7. Pattern of weekly adherence, breakdown of weekly adherence, entropy of weekly adherence and accuracy of weekly adherence prediction at the (a) minimal level and (b) full level. (Color code: blue - high adherence; orange - low adherence; purple - entropy; green – prediction accuracy)

Table 1.

Demographics of participants

Gender	Number of Participants	Percentage	Mean of Age (Year)	Std. of Age (Year)
Female	78	66%	71.5	5.1
Male	38	32%	75.0	5.8
NA	2	2%	71.0	4.2
Total	118	100%	72.6	5.5

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

Measures of attitudes and cognitive assessments

Category	Measurement	
Technology Proficiency (composite_tech_proficiency)	z_cpq	z-score for Computer Proficiency Questionnaire (CPQ) (Boot et al., 2015)
	z_mdpq	z-score for Mobile Device Proficiency Questionnaire (MDPQ) (Roque & Boot, 2018)
	z_techreadiness	z-score for Technology Readiness Questionnaire (Parasuraman & Colby, 2015)
Self Efficacy (composite_selfefficacy)	z_gse	z-score for General Self-efficacy Questionnaire (GSE) (Schwarzer & Jerusalem, 1995)
	z_tse	z-score for Technology Self-efficacy (TSE) (Schwarzer & Jerusalem, 1995)
Subjective Cognition (composite_subj_cognition)	z_iadl	z-score for Instrumental Activities of Daily Living (IADL) survey (Lawton & Brody, 1969)
	z_pdq	z-score for Perceived Deficits Questionnaire (PDQ) (Sullivan et al., 1990)
	z_mseq	z-score for Memory Self-Efficacy Questionnaire (MSEQ) (Berry et al., 1989)
Perceived Benefits (composite_perceivedbenefits)	z_indp	z-score for Brain Training and Independence Survey (Harrell et al., 2019)
	z_nict	z-score for Brain Training Belief Scale (NICT) (Rabipour & Davidson, 2015)
Objective Cognition - Reasoning (composite_reasoning)	z_ravens	z-score for Raven's Advanced Progressive Matrices (Arthur Jr & Day, 1994; Raven & Court, 1994)
	z_lettersets	z-score for Letter Sets (Ekstrom & Harman, 1976)
Objective Cognition - Processing Speed (composite_obj_cognition_processing speed)	z_ufov3	z-score for the Useful Field of View test (UFOV) (Edwards et al., 2006)
	z_digitsymb	z-score for Digit Symbol Substitution (Wechsler, D., 1997)
Objective Cognition – Memory Immediate Recall (composite_obj_cognition_ memory_immediaterecall)	z_rey_ immediate	z-score for Rey's Auditory Verbal Learning Test (AVLT), immediate recall (Schmidt, 1996)
	z_hopkins_ immediate	z-score for Hopkins Verbal Learning Test, immediate recall (Brandt & Benedict, 2001)
Objective Cognition – Memory Delayed Recall (composite_obj_cognition_ memory_delayedrecall)	z_rey_ delayed	z-score for Rey's AVLT, delayed recall (Schmidt, 1996)
	z_hopkins_ delayed	z-score for Hopkins Verbal Learning Test, delayed recall (Brandt & Benedict, 2001)

Table 3.

Details of training interactions

Category	Details
Timestamp	Date, Time
Tasks	7 Tasks (Working Memory-Updating, Switching, Tower of London, Pipe Mania, Visual Spatial, Dual N Back, Figure Weights)
Task Levels	Different number of levels for different tasks
Task Outcomes	5 outcomes (Defeat, Stalemate, Victory, Abort, Not Yet Finished)
Interaction Time	Duration in number of seconds

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4.

Correlations between the variables and the overall adherence at the minimal and full levels.

	p-values (t-test/Chi-square test)	
	Minimal Level	Full Level
Demographics		
bg_age	0.582	0.775
bg_gender	0.871	0.420
z-scores		
z_ufov3	0.435	0.417
z_digitsymb	0.836	0.537
z_ravens	0.848	0.678
z_lettersets	0.936	0.500
z_hopkins_immediate	0.182	0.207
z_hopkins_delayed	0.014	0.026
z_rey_immediate	0.070	0.076
z_rey_delayed	0.021	0.038
z_iadl	0.861	0.861
z_indp	0.863	0.649
z_nict	0.093	0.157
z_mseq	0.893	0.356
z_pdq	0.981	0.774
z_gse	0.006	0.010
z_tse	0.099	0.008
z_cpq	0.674	0.850
z_mdpq	0.557	0.939
z_techreadiness	0.087	0.045
Composite scores		
composite_tech_proficiency	0.419	0.419
composite_selfefficacy	0.006	0.001
composite_perceivedbenefits	0.405	0.259
composite_reasoning	0.574	0.302
composite_subj_cognition	0.970	0.828
composite_obj_cognition_processing speed	0.546	0.383
composite_obj_cognition_memory_immediaterecall	0.084	0.095
composite_obj_cognition_memory_delayedrecall	0.009	0.018

Table 5.

Variables derived from daily interaction details of training

Variables	Definition	Source
wd	weekday of each study day	Timestamp
task	number of tasks per day	Tasks
level	max level on each day	Task Levels
levwmu	max level on each day for game WorkingMemory-Updating	
levswi	max level on each day for game Switching	
levtol	max level on each day for game TowerOfLondon	
levppm	max level on each day for game PipeMania	
levvs	max level on each day for game VisualSpatial	
levdnb	max level on each day for game Dual N Back	
levfw	max level on each day for game FigureWeights	
def	# of Defeat on each day	
sta	# of Stalemate on each day	
vic	# of Victory on each day	
abo	# of Abort on each day	
ses	# of sessions each day	Interaction Time
sec	# of seconds each day	

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 6.

Evaluation results of overall adherence prediction models.

	Selected demographics & z-scores				Selected demographics & composite scores			
	LR	SVM	CART	RF	LR	SVM	CART	RF
Overall adherence at minimal level								
Accuracy	0.627	0.593	0.611	0.602	0.678	0.584	0.542	0.526
Precision	0.635	0.596	0.628	0.611	0.688	0.585	0.545	0.525
Recall	0.627	0.593	0.611	0.602	0.678	0.584	0.542	0.526
F1 Score	0.625	0.583	0.604	0.591	0.674	0.579	0.541	0.523
AUROC	0.672	0.645	0.613	0.648	0.708	0.655	0.543	0.591
AUPRC	0.677	0.685	0.599	0.688	0.688	0.651	0.551	0.623
Overall adherence at full level								
Accuracy	0.670	0.643	0.518	0.551	0.660	0.627	0.507	0.525
Precision	0.681	0.646	0.516	0.553	0.662	0.633	0.510	0.523
Recall	0.670	0.643	0.518	0.551	0.660	0.627	0.507	0.525
F1 Score	0.663	0.638	0.513	0.546	0.656	0.623	0.504	0.520
AUROC	0.694	0.642	0.517	0.603	0.711	0.664	0.508	0.567
AUPRC	0.677	0.655	0.525	0.610	0.688	0.677	0.517	0.615

Table 7.

Evaluation results of the weekly adherence prediction models.

	LSTM	GRU	CNN	LSTM-GRU	LSTM-Conv1D	GRU-LSTM	GRU-Conv1D	Conv1D-LSTM	Conv1D-GRU
Minimal adherence									
Accuracy	0.854	0.853	0.803	0.856	0.847	0.856	0.847	0.842	0.844
Precision	0.863	0.863	0.813	0.863	0.854	0.864	0.855	0.855	0.857
Recall	0.854	0.853	0.803	0.856	0.847	0.856	0.847	0.842	0.844
F1 Score	0.855	0.854	0.804	0.857	0.848	0.857	0.848	0.844	0.845
AUROC	0.860	0.860	0.805	0.861	0.851	0.861	0.852	0.851	0.853
AUPRC	0.735	0.734	0.666	0.740	0.727	0.739	0.726	0.719	0.721
Full adherence									
Accuracy	0.835	0.832	0.770	0.832	0.825	0.833	0.813	0.798	0.805
Precision	0.848	0.841	0.777	0.845	0.832	0.838	0.823	0.816	0.825
Recall	0.835	0.832	0.770	0.832	0.825	0.833	0.813	0.798	0.805
F1 Score	0.836	0.833	0.771	0.833	0.826	0.834	0.814	0.799	0.806
AUROC	0.844	0.837	0.771	0.840	0.829	0.835	0.819	0.807	0.817
AUPRC	0.721	0.717	0.642	0.717	0.710	0.721	0.694	0.673	0.682