



rijksuniversiteit
 groningen

Learning Companions in Neurofeedback: Do Participants Want Them?

Oisín O'Shea Kelly

Master Thesis – Clinical Neuropsychology

S6193420

October 2025

Department of Psychology

University of Groningen

Examiner/Daily supervisor: Dr.
Stephanie Enriquez-Geppert

Acknowledgements

A thesis is an aptitude test for students. The approval of the thesis is proof that the student has sufficient research and reporting skills to graduate, but does not guarantee the quality of the research and the results of the research as such, and the thesis is therefore not necessarily suitable to be used as an academic source to refer to. If you would like to know more about the research discussed in this thesis and any publications based on it, to which you could refer, please contact the supervisor mentioned.

AI Use Declaration

2. AI used for background/self-study only “I acknowledge the use of ChatGPT, 2025 to generate materials for background research and self-study in the drafting of this assessment” (Appendix F).

Abstract

Learning companions (LCs) are becoming increasingly popular for maximising the effectiveness of an individual's engagement with a brain-computer interface like neurofeedback training (NFT). The dynamic and supportive learning context which LCs are said to foster particularly important for (sub-)clinical individuals who may struggle with motivation and focus during NFT. However, before LCs become commonplace in clinical settings, it is imperative to understand the attitudes that individuals hold towards them. This study aims to assess how (sub-)clinical individuals perceive LCs over the course of three NF sessions regarding acceptance and mood. Thirteen participants (3 male, 9 female, 1 agender; $M = 24.77$, $SD = 10.47$) completed three NF sessions including an acceptance questionnaire before Session 1 and after Session 3, and the PANAS following each session. The paired samples t -test result suggests that acceptance of the LC (composite score of behavioural intention, perceived ease-of-use and perceived usefulness) significantly increased but the correlation analysis revealed that the LC did not significantly induce positive mood which led to higher acceptance scores. These findings represent an important step in the feasibility process, informing future researchers and clinicians that (sub-)clinical samples are accepting of LCs. Avenues remain however to improve their effectiveness for NFT and to further understand the association between mood and acceptance.

Introduction

Maximising the efficiency of emerging neurotechnologies is crucial for advancing our understanding of neuropsychology and for helping those who are afflicted with various mental disorders.

Neurofeedback training (NFT) is an example of self-neuromodulation via brain-computer interface (BCI) which has become increasingly popular as a way to improve symptoms of depression (Fernandez-Alvarez et al., 2022), anxiety (Tolin et al., 2020) or ADHD (Vlachou et al., 2022) while also helping to improve abilities such as memory (Jackson et al., 2023) or even athletic performance (Brito et al., 2022). NFT requires individuals to participate in a number of sessions with a specialist, whereby a curated protocol helps participants to self-modulate their brain activity as real-time feedback is provided via electroencephalogram (EEG). Despite many promising results, there are still many individuals who do not effectively engage with NFT (Kadosh & Staunton, 2019; Loriette et al., 2021). This occurrence of non-responders has led researchers to strategise on what are the underlying variables which prevent effective engagement with NFT, and what strategies may resolve this.

Learning companions (LCs) have become more popular over recent years in an effort to improve the engagement and efficiency of user interaction with BCIs. LCs can come in various forms and are often closely connected with other terms such as social robot or pedagogical agent. Specifically however, one of the first definitions from Chou et al. (2003) described an LC as a non-human computer-based intelligent tutoring system which fosters learning through social interaction but does not hold an authoritarian influence towards the individual. LCs often take a partly or wholly digital form with warm, anthropomorphic features such as big, blinking eyes or a wide smile (Song et al., 2021). LCs interact with users in a number of ways. They act as co-learners, explaining key terms if necessary or they promote specific learning strategies by encouraging users to reason aloud or ask questions. LCs also provide corrections in a supportive way (as a peer rather than a superior) and their social presence makes the task of learning feel more interactive. Thus, LCs enrich the learning context, causing it to become more dynamic and interactive (Chou et al., 2003). The positive emotions which this is said to induce is one of the key mechanisms behind the improvements that LCs seem to cause (Pillette et al., 2020; Han et al., 2025). Task engagement (Zielke et al., 2024), flexibility and

creativity (Earle-Randall et al., 2024), motivation (Lester et al., 1999), and task efficiency (Kim et al., 2006) are all variables which LCs are said to enhance. Such improvements have been noticed in a number of different contexts. Classrooms and other educational settings tend to be the most common locations for LCs, but other studies demonstrate their effectiveness in medical or military contexts (Johnson & Lester, 2018; Stommel & Stommel, 2021). Most relevant to this study, Pillette et al. (2020) demonstrated that for those who struggle with BCI usage, LCs represent a viable avenue to enhance their ability to learn and memorise how to use the BCI while also improving the subjective user experience. These results lay a promising foundation for the prospect of improving neurofeedback results with LCs. However, given that LCs are a relatively recent technology, understanding the attitudes of individuals towards them is key to maximising their efficiency.

A user's attitudes and beliefs towards a tool such as an LC is likely to influence their success in using it and to predict overall adoption of the technology. Research suggests that when individuals are more accepting of an LC, they demonstrate higher enthusiasm and fewer mistakes during usage (Kort et al., 2001). Previously, this concept had been discussed via the Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Usage of Technology (UTAUT). Those theories and models aimed to discover the variables which are most important when considering usage and acceptance of social robots or LCs. Such variables include usefulness, adaptability, enjoyment, sociability, companionship and perceived behavioural control (Venkatesh et al., 2003; Davis, 1989). More recently however, Grevet et al. (2024) have expanded the concept of user acceptance to publish a BCI acceptability questionnaire. This work identifies a number of important latent traits with much focus being placed upon the target variable of behavioural intention (BI) with its moderators of perceived ease of use (PEOU) and perceived usefulness (PU) also being of great importance for acceptance. BI is a measure of an individual's intention to use the tool/technology. If BI is high, then the user is more likely to adopt the tool. PEOU refers to the thoughts a user may have regarding how simple the tool will be to interact with. If the PEOU is high, then the user believes that interacting with the tool will be relatively free of effort. PU instead refers to thoughts the user may have pertaining to how useful the tool will be to them. If the PU is high, then the user believes that the tool

will help them to achieve their desired outcome. PEOU and PU represent the most important predictors of BI. Although BI represents a main outcome variable within the work from Grevet et al. (2024), incorporating other factors such as PEOU and PU within the outcome variable may result in a measure of an individual's acceptance which offer further insight into how easy they find the LC to interact with and how relevant they find it in an NFT context. With this in mind, future references to acceptability/acceptance in this work will denote a composite score of BI, PEOU, and PU (further detail on the validity of the measure is described in the *Data Analysis* section of the *Methods*). Notably, it is also important to distinguish between concepts of acceptability and acceptance. Acceptability refers to the attitudes and intentions a user may have *prior* to interacting with or using the tool/technology (Alexandre et al., 2018). Acceptance on the other hand refers to the evaluation of the tool/technology *after* interaction or usage (De Graaf & Allouch, 2013). These concepts are important to consider for all individuals when a practice involves a tool such as an LC, but it is even more so when working with clinical or subclinical populations. For example, symptoms of anxiety and depression have had a notable detrimental effect on BCI usage in previous studies (Jeunet et al., 2016). Understanding this relationship between the individual and the LC they interact with informs researchers and clinicians as to how necessary and impactful they can be.

The affective component which LCs offer is one of the main appeals of the technology. Specifically, this affective component can be seen in studies where participants exhibit both reduced negative affect feelings [stress, anxiety etc] and higher positive affect feelings [motivation, enjoyment etc] (Han et al., 2025; Edwards et al., 2020). Emotional Response Theory is one purported mechanism behind this advantage with the verbal (expressing interest, offering feedback/support) and non-verbal cues of the LC (smiling, dynamic vocal tones) being of benefit to the mood of the individuals (Liew et al., 2017). Despite these advantages, it is not a guarantee that the mood of an individual will be improved following usage of an LC. According to Cognitive Load Theory, the additional cognitive processing of an LC could be of detriment to individuals who are usually already undertaking something which requires focus (Liew et al., 2017). This is of particular concern with a procedure like NFT which applies significant mental load on the individual and which requires significant mental

resources (Bauer & Gharabaghi, 2015). Furthermore, scepticism has been voiced on the topic of introducing social robots or LCs to the general public. Concern has been raised regarding the coercion of vulnerable or less independent populations into using such a technology when they would rather be with a human (Share & Pender, 2018; Morris, 2021). If the presence of an LC is something which hinders the mood of an individual, their necessity in both learning and other contexts is called into question. The importance of understanding how mood may be affected by an LC becomes even more important when considering prior research which posits mood as a predictor of BCI performance in some users (Nijboer et al., 2010). Thus, it remains imperative to listen to those who interact with LCs (particularly when they represent vulnerable populations such as sub-clinical/clinical individuals) to avoid coercion and to ensure that the presence of such a technology is not detrimental to their mood.

Feasibility studies represent a way to learn more about how an intervention can be evaluated and implemented to inform researchers and clinicians if it is worth the costs and resources, but also to identify what are the strong and weak points which can be addressed. They help to improve internal and external validity through the identification and removal of methodological issues or through gaining a better understanding of the resources necessary for its implementation (Fredericks et al., 2019). Such improvements have already been found in feasibility studies which explore BCI technology ranging from setup procedures becoming more participant friendly (Mansour et al., 2025), to resource allocation becoming better understood (Lim et al., 2023), to better insights into acceptance being gained (Grevet et al., 2023). Thus, feasibility studies ask the transdiagnostic question of “can this be done?” (National Institute for Health Research, 2012). Undertaking a feasibility study is greatly beneficial for exploring the concepts of acceptability, acceptance and mood (Gadke et al., 2021), factors which are imperative when looking to combine two modern technologies such as LCs and NFT.

The present study aims to investigate the influence that interacting with an LC during frontal-midline theta NFT has on individuals’ acceptance of the technology. To this end, participants’ scores of the acceptability questionnaire (Grevet et al., 2024) which was completed before the first and after the third NFT session will be compared.

Assessing the influence of interaction with an LC on participants' mood is another goal of this study. This will be achieved by analysing the results of the PANAS questionnaires which were administered to participants following each of the three NFT sessions. Thus, the completion of this study should provide a broader view of the way in which the usage of LCs during NFT impacts the mood of (sub-)clinical individuals and how this affects acceptance. Specifically, the following hypotheses are proposed:

Hypothesis 1, H1: using the LC during NFT will result in participants' acceptance of it being significantly higher than the acceptability.

Hypothesis 2, H0: using the LC during NFT will not result in participants' acceptance of it being significantly higher than the acceptability.

Hypothesis 3, H1: interaction with the LC induces positive feelings which result in high scores on the PANAS scale and also in high acceptance scores.

Hypothesis 4, H0: interaction with the LC does not significantly induce positive feelings and thus does not result in high scores on the PANAS scale and/or contribute to high acceptance scores.

Methods

Ethical Procedure, Recruitment and Participants

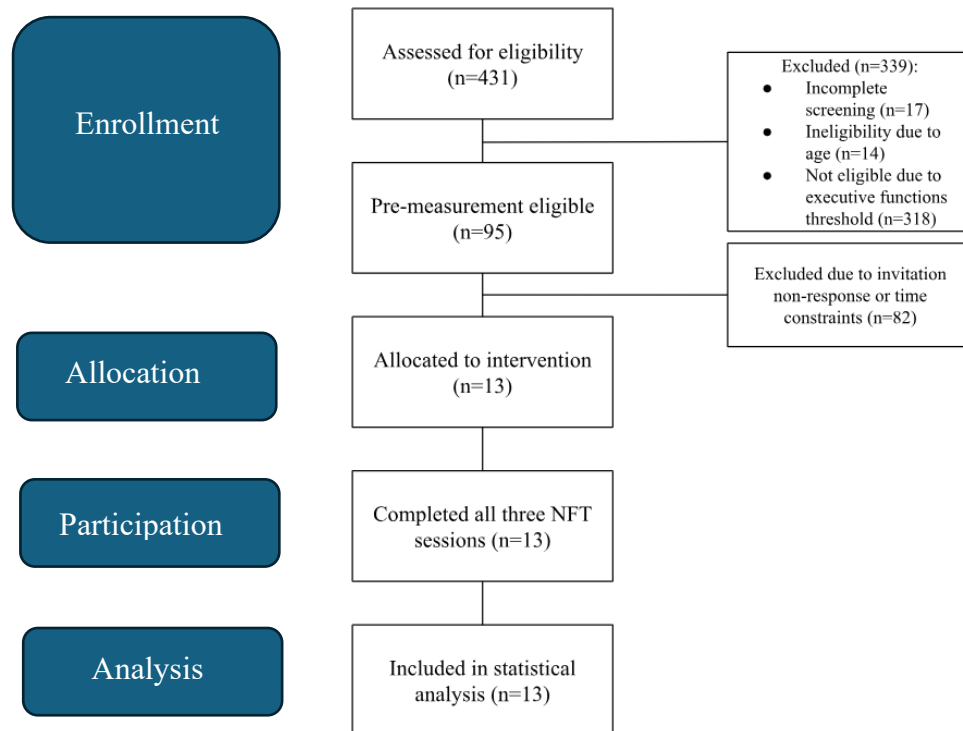
Ethical approval for this study was granted by the Institutional Review Board, documented under the ethical reference number PSY-2324-S-0092.

Participants were recruited via the SONA credit system (mandatory for first-year psychology students), via posters which were hung up around University of Groningen faculties, and via the researchers' personal networks. Participants completed the Behaviour Rating Inventory of Executive Function - Adult Version (BRIEF-A) to assess self-reported everyday EF impairments.

Inclusionary criteria necessitated participants being 18 or above years of age and scoring in the 75th percentile or higher in the Behaviour Rating Inventory Executive Function-Adult version (BRIEF-A; Roth et al., 2005) or on any of its subscales. Furthermore, the presence of a brain tumour or epilepsy excluded participation, as did colour-blindedness and/or the usage of a psychoactive drug (cannabis, alcohol, nicotine and caffeine excluded) in the previous three weeks of the study. Based on these criteria, 95 individuals were eligible and 13 took part in the study.

The age of participants ranged from 18-54 ($M = 24.77$, $SD = 10.47$). The sample consisted of 3 males, 9 females and 1 who described themselves as agender. Pertaining to the presence of a psychiatric or neurological disorder, 4 were officially diagnosed, 3 were not officially diagnosed but suspected one or more may have been present, and 6 reported that none were present. Generalised anxiety disorder (GAD) and attention-deficit hyperactive disorder (ADHD) were the most common disorders with a more extensive table detailing the disorders below (Appendix A).

Figure 1. *CONSORT Flowchart of Recruitment Process*



Design

A single group pre-post design without a control group was used throughout the study. The design aligns with the TULIP project at RUG, which is a feasibility study. The goal was to assess the acceptability of the LC during NFT with participants with executive function (EF) impairments rather than to study the intervention's effectiveness, thus a control group was not necessary.

Procedure

Information and informed consent

Participants who wished to partake in the study received an information sheet and an informed consent sheet. The information sheet included a description of the purpose of the research, the screening procedure, the right of withdrawal for participants, information regarding data treatment and storage, some benefits and possible risks and also the email addresses of the researchers in case participants had further questions. After the participants had read the information sheet, an informed consent sheet was presented. Through signing the informed consent, the participant agreed with all of the information given in the information sheet.

Neurofeedback procedure

All of the data collection occurred in a sound-attenuated EEG lab located in the Heymans institute at the University of Groningen. All participants followed the same protocol: three neurofeedback sessions scheduled involving the participant and two members of the research team. It was aimed to schedule the three sessions within a single week and at the same time of day, in order to maintain regularity and avoid distractions between, for example, morning and afternoon sessions.

Sessions took approximately two hours for the participant with another 40 minutes (approximately) of preparation and cleaning down tasks for the researchers present. Upon arriving for the first neurofeedback session, the participant completed three questionnaires: the pre-assessment questionnaire, the acceptability questionnaire, and the personality questionnaire. Then, the EEG cap setting and calibration was undertaken, including a blink-threshold procedure so blinks could be filtered out without losing valid data. After a practice block and a baseline pre-block, six NF blocks commenced. Mental strategies aimed at upregulating FM theta activity were attempted by participants during these blocks as they received immediate feedback on the computer screen in the form of colour-coded squares (green = upregulation, red = downregulation, grey = artifact). Following each block, two sentences of feedback were given by the LC and the logbook was filled out. The NF session ended with a post-rest block and the completion of the PANAS questionnaire.

The second and third sessions followed the same procedure, but no questionnaires were completed before training. After the third session, participants filled in the PANAS, the Companion Usability questionnaire, the Acceptability questionnaire, and the BRIEF-A. Data was securely stored on the university server. SONA participants were granted their course credits; beyond this no further compensation was offered.

Materials

Juno - Learning Companion

Juno was the name of the robot-like learning companion involved in this research. Juno was developed by the TULIP project group (Enriquez, 2024). Juno featured interactive eyes via a

smartphone app placed at the front of the head of the styrofoam design (Appendix B). Feedback was provided by Juno after each NF block by delivering one statement of encouragement (e.g support effort, general effort) and one strategy-related suggestion (strategy keep or strategy change). Examples of encouragement included: “You have made a remarkable effort” (support effort) or “Have patience, you will progress” (general effort). Examples of suggestive feedback included: “Maintain the mental strategies that have been successful” (strategy keep) or “Consider exploring a different mental strategy” (strategy change). The LC feedback was offered exclusively in English as this was the only language which had been prepared by the time of data collection.

Participant Logbook

Participants were asked to complete a logbook following each block. The mental strategies they engaged in were filled in and rated on a scale from 1 to 7 (1 = not effective, 7 = very effective). Participants could record up to 8 strategies which they may have used.

Mental Strategies List

A list of mental strategies in Dutch and English was provided during NF blocks. The list offered suggestions for which strategies to engage in rather than enforcing participants to use those particular strategies. The list (Appendix C) was designed by researchers in the TULIP research project.

Questionnaires

Pre-assessment questionnaire: This gathered demographic information (age, gender, occupation etc), self-report data on psychological/neurological/psychiatric conditions, and it included the BRIEF-A.

Acceptability questionnaire: This questionnaire derives from the BCI/Neurofeedback Acceptability Tool (Grevet et al., 2023). This tool was developed based on previous models such as the Technology Acceptance Model 3, the Unified Theory of Acceptance and Use of Technology (UTAUT), and the components of user experience (CUE) model. The adapted version used in this study consists of 50 questions which examine 18 variables (1-3 questions per variable). Previous research deemed BI the most relevant variable for overall use behaviour with PEOU and PU being key predictors of BI. Four

categories contributed to each of the three above variables. These categories include system characteristics, facilitating conditions, social influence and individual differences. Each of these categories were made up of multiple questions which assessed further specific subcategories (Grevet et al., 2023). All questions in the tool use an analogous scale where participants can provide any answer ranging from 0 (“totally disagree”) to 100 (“totally agree”) in response to a statement such as “If I had the opportunity, I would like to use the neurofeedback companion again during my neurofeedback training for the improvement of cognitive abilities”. This tool demonstrates promising psychometric properties with Cronbach’s alpha ranging from .83 to .97 for perceived ease-of-use, perceived usefulness, and behavioural intention, suggesting good internal consistency. Furthermore, a good fit between the model and the dataset providing support for its validity and utility is apparent with a value of .913 for comparative fit index and .897 for Tucker-Lewis Index.

Personality questionnaire (16PF-5): This measured 16 primary traits and 5 higher order (global) factors. Internal consistency is typically $\alpha \approx .75$; test-retest reliability $r \approx .70-.80$ (primary scales) and $r \approx .78-.87$ (global factors). The 16-factor structure is supported, with good convergence with the NEO-PI-R (administered before session one; Cattell & Mead, 2008)

BRIEF-A: Assesses everyday EF difficulties; internal consistency is excellent ($\alpha \approx .94-.96$) and validity has been demonstrated across clinical and non-clinical samples (used in the pre-assessment and re-administered after session three; Roth et al., 2013).

PANAS: Assesses the emotional status of participants following each of the three neurofeedback sessions. Two 10-item scales measure positive and negative affect respectively. An example of a question which measures positive affect is “*Indicate the extent you have felt this way over this session. - Enthusiastic*”, to which the participant provides an answer on a 5-point Likert scale ranging from “*very slightly or not at all*” to “*Extremely*”. An example of a question which measures negative affect is “*Indicate the way you have felt this way over this session. - Ashamed*”. PANAS has a strong internal consistency reliability (PA $\alpha \approx .86-.90$; NA $\alpha \approx .84-.87$) with good test-retest values (up to $r \approx .71$) and a clear two-factor structure with low intercorrelation supporting discriminant validity (Watson et al., 1988; Heubeck and Wilkinson, 2019).

Companion Usability: Adapted from Pilette et al., (2019) comprising four dimensions:

Learnability/Memorability, Efficiency/Effectiveness, Safety, and Satisfaction. Items were rated on a 1–5 Likert scale (administered after session three).

Data Analysis

EEG data

Data extraction was done per block per person, processed in MATLAB and normalized to 1 to 30 hertz (i.e. mean theta amplitude / mean amplitude in the full frequency band).

Data preparation

A composite acceptability/acceptance score was derived by averaging the means of BI, PEOU and PU. This decision was supported by other studies which examine acceptance via multiple variables rather than just BI (Rosli et al., 2022), and by the motivation to have an acceptance variable which fully incorporates how participants felt regarding the LCs specific NFT effectiveness, and how easy they found interaction with the LC in this context. Internal consistency for each subscale (BI, PEOU, and PU) and for the composite acceptability/acceptance variable was assessed using Cronbach's alpha. Inter-subscale correlations were also examined to assess whether the subscales measured related aspects of a common construct. A full breakdown of these statistics is provided in Appendix D.

A number of variables were also highlighted as warranting respective exploratory analyses based off the predictive value that prior literature suggests they may hold for acceptance (Grevet et al., 2024). These included BI, PU and PEOU respectively. Furthermore, the categories of system characteristics (SC), social influence (SI), facilitating conditions (FC) and individual differences (ID) were also assessed to gain the best possible understanding of the factors which contribute to an individual's acceptance of an LC.

Data preparation involved testing for outliers and assumptions of normality for the paired samples *t*-test. Due to the small sample size, the mean absolute deviation method was used to test for outliers (Leys et al., 2013). No values were excluded due to this method. Shapiro-Wilk tests and q-q

plots were used to check for assumptions of normality for every paired samples *t*-test which was undertaken. All of the data complied with the assumptions and so paired samples *t*-tests were chosen to assess the differences. Holm-Bonferroni corrections were applied to the tests where necessary.

The total positive affect (PA) and total negative affect (NA) scores from the PANAS were used for the correlation analysis. The mean of each of these scores throughout each of the three sessions was found so that a measure of the participants' average mood after interacting with the LC could be used. Acceptance was tested against the mean PA and the mean NA scores to assess the strength of the correlation.

Before conducting the correlation analysis, assumptions of linearity, normality, and homoscedasticity were checked along with checking for the presence of outliers. Again, no outliers were found here with the mean absolute deviation method being used to check.

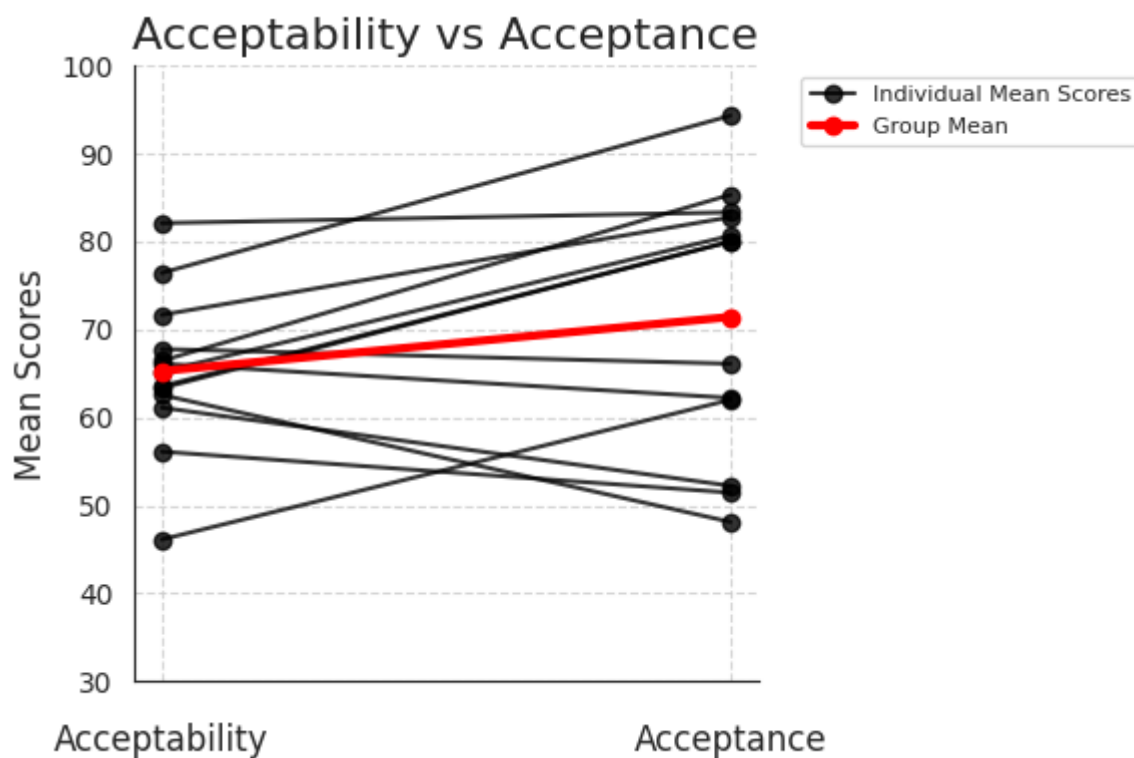
Non-parametric alternatives were conducted alongside the planned parametric tests for both research questions. This was done even when all of the assumptions of the parametric tests were met. The purpose of this was purely to enhance the statistical rigor of the results which may have been initially questioned due to the small sample size. The results of the non-parametric alternatives are supplementary and do not directly contribute to the interpretation of the hypotheses.

Bootstrapping resampling analyses (10,000 samples) were also conducted to obtain an empirical confidence interval for the mean differences in order to further strengthen the robustness of results which may have been initially questioned due to the small sample size.

Results

A paired samples *t*-test was conducted to assess if interaction with the LC over three sessions led acceptance scores to be significantly higher than acceptability scores. Results indicated that the mean score for acceptance ($M = 71.43$, $SD = 15.1$) was significantly higher than that for acceptability ($M = 65.27$, $SD = 8.85$). The difference, 6.15, 95%[0.31, inf], was significant $t(12) = 1.88$, and $p = 0.043$ and represented a moderate effect, $d = 0.5$.

Figure 2. Comparison of Acceptability (Before Session 1) and Acceptance (After Session 3) Mean Scores.



Note. Y-axis does not begin at 0.

A supplementary nonparametric Wilcoxon signed rank test was also performed. The result of this test¹ supported that of the above *t*-test. A supplementary bootstrap analysis (10,000 samples) was also

¹Acceptance scores ($Mdn = 80 = IQR = 20.78$) were significantly higher than acceptability scores ($Mdn = 65.22$, $IQR = 5.22$), $W = 70$, $n = 13$, $p = .047$, $r = 0.54$.

conducted to assess the robustness of the finding. The bootstrapped 95% one-sided lower confidence limit (0.88) was above zero, suggesting a precise estimate of effect size.

Exploratory, one-tailed paired samples *t*-tests were conducted to assess if interaction with the LC led the variables of BI, PEOU, and PU to increase over the three sessions.

On average, interaction with the LC did not lead BI scores to be higher post-interaction ($M = 70.39$, $SD = 21.3$) compared to pre-interaction ($M = 69.28$, $SD = 12.92$). The difference, 1.1, 95% [-7.74, inf], was not significant $t(12) = 0.22$, and $p = 0.68^*$ and represented a small effect, $d = 0.06$. A supplementary nonparametric wilcoxon signed-rank test was also performed² which supported the findings of the *t*-test. A supplementary bootstrap analysis (10,000 samples) was also conducted, finding that the 95% one-sided lower confidence limit (-6.64) was below zero, indicating that the effect may not be statistically reliable.

On average, interaction with the learning companion did not lead PU scores to be higher post-interaction ($M = 68.05$, $SD = 19.86$) compared to pre-interaction ($M = 65.95$, $SD = 10.61$). The difference, 2.1, 95% [-6.77, inf], was not significant $t(12) = 0.42$, and $p = 0.68^*$ and represented a small effect, $d = 0.13$. A supplementary nonparametric wilcoxon signed-rank test was also performed³ which supported the findings of the above paired samples *t*-test. A supplementary bootstrap analysis (10,000 samples) was also conducted, finding that the bootstrapped 95% one-sided lower confidence limit (-5.82) was below zero, indicating that the effect may not be statistically reliable.

On average, interaction with the learning companion led PEOU scores to be higher post-interaction ($M = 75.85$, $SD = 12.06$) compared to pre-interaction ($M = 60.64$, $SD = 11.74$). The difference, 15.21, 95% [8.67, inf], was significant $t(12) = 4.15$, and $p = 0.002^*$ and represented a large effect, $d = 1.28$. A supplementary nonparametric wilcoxon signed-rank test was also performed⁴ which supported the findings of the above paired samples *t*-test. A supplementary bootstrap analysis (10,000 samples) was also conducted, finding that the bootstrapped 95% one-sided lower confidence limit (9.54) was above zero, indicating that the effect is statistically reliable.

²Post BI scores ($Mdn = 72.33$, $IQR = 36.67$) were not significantly higher than pre BI scores ($Mdn = 67$, $IQR = 20$), $W = 47$, $p = 0.473$, $r_{bc} = 0.03$

³Post PU scores ($Mdn = 66.67$, $IQR = 12.33$) were not significantly higher than pre PU scores ($Mdn = 66.67$, $IQR = 25$), $W = 52$, $p = 0.17$, $r_{bc} = 0.33$

⁴Post PEOU scores ($Mdn = 80$, $IQR = 13.67$) were significantly higher than pre PEOU scores ($Mdn = 62$, $IQR = 13$), $W = 87$, $p < 0.001$, $r_{bc} = 0.91$

* = value corrected via Holm-Bonferroni method

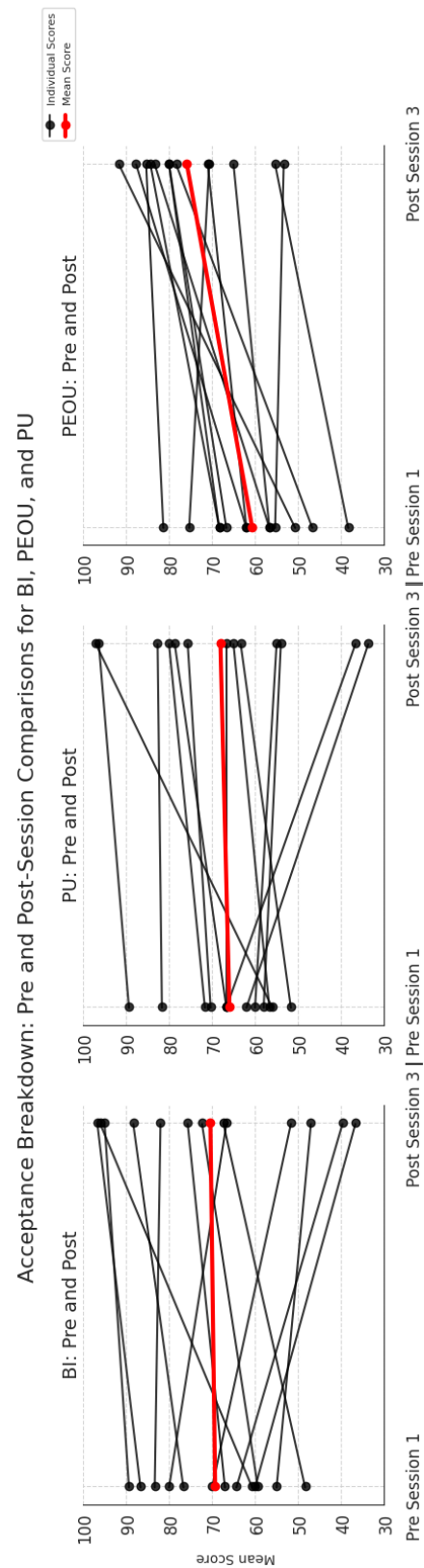
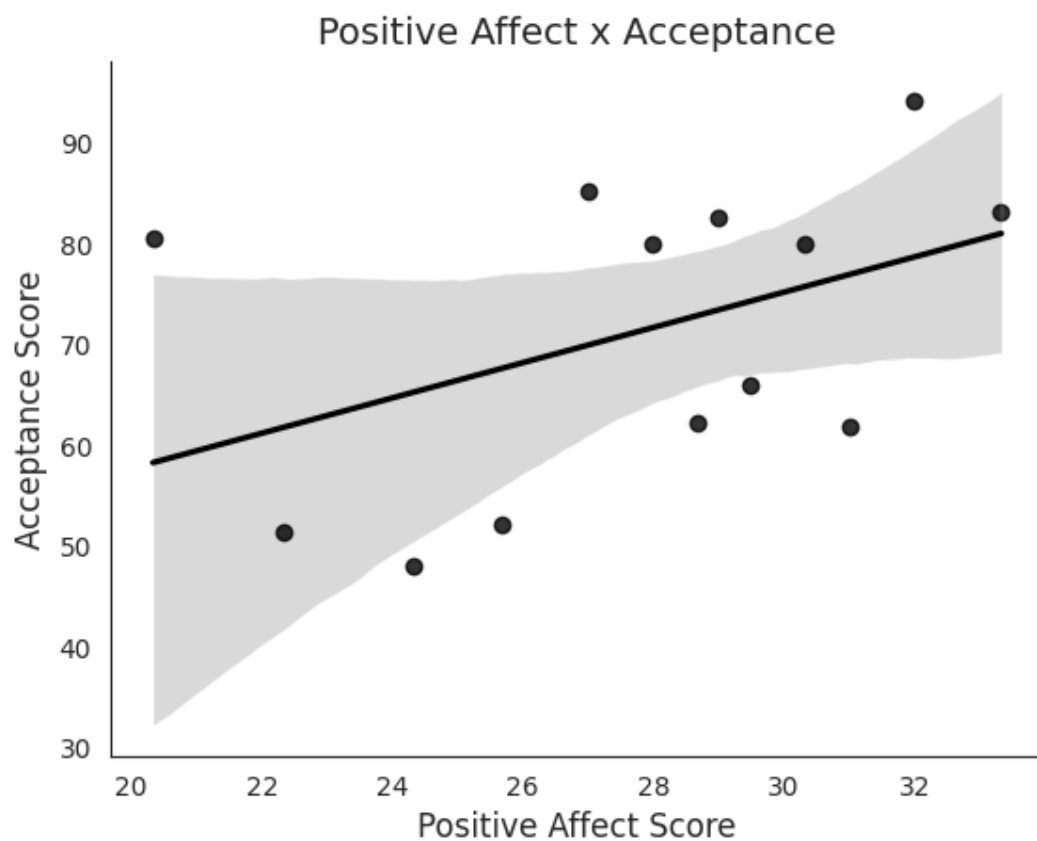


Figure 3. Respective Comparison of Mean Scores of BI, PU, and PEOU Before Session 1 and After Session 3. Red Line Indicates Overall Mean Group Score. *Note:* Y-axes do not begin at 0.

Four exploratory paired-samples t-tests were conducted to assess the categories of system characteristics, facilitating conditions, social influence and individual differences. None of these tests indicated a significant change in the variables after interaction with the LC following Holm-Bonferroni correction (see Appendix E for full breakdown).

A Pearson's correlation analysis was conducted to test H2. The results showed a non-significant, moderate positive relationship between positive affect of participants' and their acceptance of the learning companion following their third session, $r(11) = 0.44$, $p = 0.065$, 95% CI[-0.05, 1.0].

Figure 4. Correlation Analysis Between Positive Affect and Acceptance. Grey Shading Represents 95% Confidence Intervals.



Note. Axes do not begin at 0.

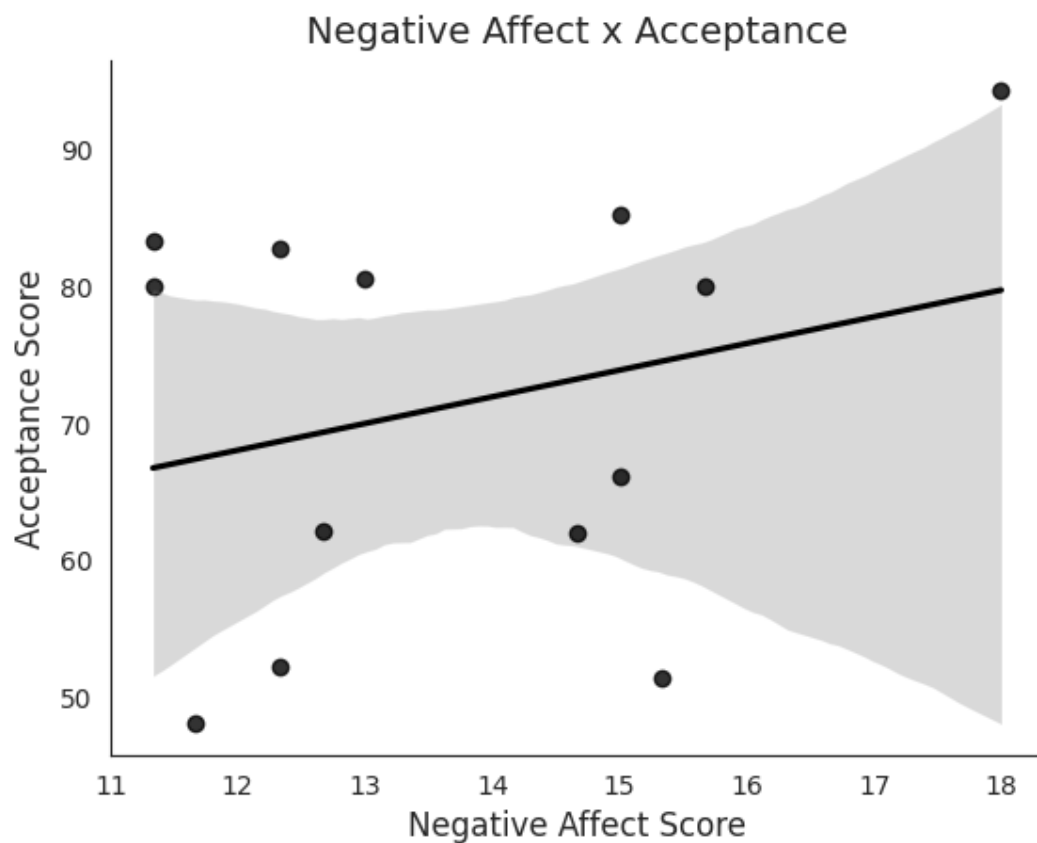
This relationship was also assessed via a non-parametric Spearman's rank-order correlation, the results of which⁴ supported the parametric correlation findings. A supplementary bootstrap analysis

⁴ $r_s(11) = 0.457$, $p = 0.058$

(10,000 samples) was also conducted, finding that the bootstrapped 95% one-sided lower confidence limit (-0.008) was below zero, indicating that the effect may not be statistically reliable.

An exploratory analysis was conducted to assess the relationship between negative affect of participants' and their acceptance of the learning companion following their third session. This analysis was not directional. The findings indicated a small, non-significant positive relationship, $r(11) = 0.26, p = 0.384, 95\% \text{ CI}[-0.34, 0.71]$.

Figure 5. Correlation Analysis Between Negative Affect and Acceptance. Grey Shading Represents 95% Confidence Intervals.



Note. Axes do not begin at 0.

This relationship was also assessed via a non-parametric Spearman's rank-order correlation, the results of which⁵ supported the parametric correlation findings. A supplementary bootstrap analysis

⁵ $r_s(11) = 0.16, p = 0.299$

(10,000 samples) was also conducted, finding that the bootstrapped 95% confidence interval [-0.49, 0.727] included zero, indicating that the effect may not be statistically reliable.

Discussion

Present findings

This study endeavoured to further the current understanding of how engaging with an LC makes (sub-)clinical individuals feel in an NFT setting. The results suggest that the interaction between participants and the LC over the three NF sessions caused them to become significantly more accepting of it. The results did not however indicate that there were strong positive emotions induced which increased acceptance rates. The implications of these findings are discussed below.

Acceptance

When comparing acceptability to acceptance scores, the results showed that participants became significantly more accepting of Juno over the course of the three sessions. This finding from the paired samples *t*-test was also supported by the supplementary nonparametric equivalent and the bootstrapping analysis, highlighting the robustness of the results despite the small sample size. This finding is consistent with the work of Pillette et al. (2020), strengthening the argument for LCs to be involved in BCI procedures. To truly understand however why this significant increase with a moderate effect occurred, it is important to examine the three variables which contributed towards acceptability/acceptance within this study.

Behavioural Intention

BI of participants towards the LC (i.e their intention to use Juno in future similar situations) did not increase significantly following the three sessions. However, there was an overall mean increase in scores, and the pre-interaction scores were already relatively high compared to many other variables. Participants, when confronted with questions like “*Assuming I had access to a neurofeedback companion again during my neurofeedback training, I would use it*”, responded generally positively. Although the overall mean increase and relatively high pre-interaction scores suggest that BI is a variable which does not elicit immediate concern for future researchers/clinicians, the nonsignificant increase underlines opportunity for improvement. Social influence (SI), another category within the BCI-Acceptance questionnaire is a known predictor of BI (Grevet et al., 2023). SI results were also

not significant in this study (Appendix E). Previous literature suggests that showing excerpts or short quotes from previous participants that are positive should lead to higher SI scores (Chao, 2019; Cao et al., 2024). Doing this in future applications of LCs could indirectly improve BI and lead participants to show a stronger likelihood to be willing to adopt a technology (such as an LC) over time.

Perceived Ease of Use

Interaction with the LC led PEOU scores to increase significantly across the three sessions. Thus, participants answered questions like *“I think practicing neurofeedback with the neurofeedback companion is easy”* much more positively after they had interacted with Juno. The strong effect observed here highlights the success of the design of the LC in this study. Widening eyes, raising eyebrows, physical form and voice are all aspects of an LC which contribute to participants finding the interaction to be comfortable and free of effort (Kapoor et al., 2001; Belpaeme et al., 2018; Schreibelmayer & Mara, 2022). The design of the LC in this study was carefully informed by prior research (Pillette et al., 2020) to include such features, an effort which results indicated has paid off. Some previous research has expressed concern about how an LC could provide a learner with unnecessary cognitive load during an already cognitively expensive task such as NFT (Liew et al., 2017; Bauer & Gharabaghi, 2015). This was considered particularly alarming for (sub-)clinical samples (Barth et al., 2021). However, our findings detract credibility from this idea. Instead, the results suggest that sufficient detail given to the design of an LC makes participants likely to find it easy to interact with and to cognitively process, and thus suitable even for complicated tasks like NFT.

Perceived Usefulness

As PU results indicate, participants did not respond to prompts such as *“In my opinion, adding a neurofeedback companion is useful in the context of neurofeedback for the improvement of cognitive abilities”* significantly more positively after interaction with Juno. There again was an overall mean increase in PU results after session three compared to session one, but this increase was very minor. Such a result is important in a feasibility context. It highlights an area which should not necessarily provoke discussion about how unsuitable an LC is during NFT, but that does require improvement to

maximise the efficiency of the technology. Previous findings have cited system characteristics (SC) as a category which influences PU (Grevet et al., 2024). Within our analysis, this category also did not show significant increases over the three NFT sessions (see Appendix E). Targeting this category and its respective variables could represent a viable way to increase PU. For example, although the researchers in this study did briefly explain the purpose of Juno to the participants who also completed pre-session questionnaires which introduced the concept of an LC, it is possible that the specific purpose the LC was not fully understood. NFT represents a modern technology that many individuals are vastly unfamiliar with (Eisenbarth et al., 2025). Coupling this with another modern (and likely unfamiliar) technology such as an LC might necessitate more in-depth instructions from researchers regarding the mechanisms at play. Previous studies have highlighted the instructional style of researchers as imperative in BCI work (Lotte et al., 2013). An extra layer of detail provided to the participants in this study surrounding the specific mechanisms which make Juno important (e.g how the real time data informed Juno's feedback and how this helps to increase EF over the sessions) may have helped to assure the participants' as to the relevance of the LC and thus to increase their PU.

Positive Affect x BI

A correlation analysis conducted between the acceptance scores and the positive affect score from the PANAS, revealed a moderate, positive, non-significant relationship. Although the effect size is a somewhat promising finding, the non-significant result implies that there was not a strong enough presence of positive emotions (self-confidence, enjoyment, motivation) within participants to make them more inclined to have positive attitudes towards the LC. Two possibilities may have contributed to such a result: (1) positive emotions were not effectively induced by the LC, or (2) the positive emotions induced had little impact on participants' acceptance scores.

The first possibility is somewhat supported by the exploratory analyses relating to the first research question. Specifically, the facilitating conditions category was subject to only a very minor mean increase across sessions and this increase was not significant (Appendix E). Facilitating conditions is a variable within the BCI-Acceptability questionnaire which encompasses factors such as “playfulness”, “ease of learning”, and “social/emotional support”. Had the LC inspired the positive

affect among participants that was expected given prior research (Han et al., 2025), this category would likely have seen a significant increase across sessions and a higher mean score than was actually achieved. One potential shortcoming of Juno in this regard, is that the feedback provided was limited. Following each NF block, Juno provided two sentences to update the participant on their progress and how it may be improved. If one particular block progressed similarly to a previous one, the same piece of feedback could be repeated to the participant. Research indicates that more extensive feedback that varied more between sessions may help to inspire feelings of motivation or social support (Ortiz-Ordíñez et al., 2015; Pilette et al., 2020).

It remains plausible that the second aforementioned possibility remains true and that positive emotions were significantly induced yet they just had a weaker than expected impact on acceptance. This idea is supported by the generally high PA scores found in this study compared to others that used PANAS (Thompson, 2007). The weaker than expected impact on acceptance may be particularly relevant for our sample of (sub-)clinical participants. Although previous literature is clear on the importance of such affective factors regarding acceptance (De Graaf & Allouch, 2013), the (sub-)clinical sample within this study may have placed less emphasis on such factors compared to a sample without the same symptoms. For example, someone undertaking NFT to reduce the functional impact of ADHD related symptoms may be more goal-oriented than someone without such symptoms (Hasslinger et al., 2020). Therefore, this (sub-)clinical sample may have been less impacted by affective components of the LC by being more focused on its functional impact.

Despite the correlation between positive affect and acceptance not being as strong as initially hypothesised, the results of this research question are more promising when one considers that this is a feasibility rather than an intervention study. The moderate positive correlation doesn't rule out the possibility that future similar studies should target positive affect as a means to increase the acceptability of an LC for NFT, especially given that the sample size was relatively small, and p -values were only slightly above the significance thresholds. Additionally, the high mean positive affect score compared to previous mean scores from studies which examined PANAS (Thompson, 2007) indicates that this result does not raise significant concern regarding how an LC may impact the

mood of a (sub-)clinical participant during NFT. This implication is also furthered by the exploratory result of this research question.

Negative Affect x BI

An additional correlation analysis was conducted with the mean negative affect scores of participants after each session and their acceptance scores. This analysis exhibited a small, non-significant positive correlation, thus indicating that the negative affect that participants felt following the sessions had very little bearing on their acceptance of the LC. This finding, coupled with the observation that the NA scores were not significantly higher than what would have been expected based on prior literature (Thompson, 2007), is promising for future research. It further weakens the argument from Cognitive Load Theory, that the additional processing of an LC will burden the mind and hinder the mood of a participant who is already exerting significant cognitive effort during the session (Liew et al., 2017).

Hence, the findings from this research question inform future researchers or clinicians more so on the lack alarming trends regarding how LCs impact the mood of (sub-)clinical individuals during NFT, rather than the strength of the association between mood and acceptance.

Limitations

The findings of this study should be interpreted alongside its limitations. Regarding the first research question, the hypothesis hinges on the changes in the acceptability/acceptability variable. Although the decision to utilise a composite acceptability/acceptance variable derived from BI, PEOU, and PU attempted to provide a broad overview of acceptance variables, tests examining the inter-item correlations between these three factors in this study highlighted inconsistencies. Specifically, calculations of internal consistency for PEOU were considerably lower than standard thresholds (Tavakol & Dennick, 2011) and also than both BI and PU (see Appendix D). Although PEOU was not removed from the scales contributing to acceptability/acceptance, interpretations of the outcome variable should consider this finding.

Furthermore, methodological limitations arise when reflecting upon the sample size and number of sessions involved within this study. Thirteen participants were involved in three NFTs.

Although this is a feasibility study rather than a full-scale implementation of a procedure/protocol, the limited sample size hinders the generalisability of the findings. Additionally, the limited number of actual NFTs (just three) and thus exposure to the LC means that we may be viewing more of a brief snapshot of attitudes towards the technology, rather than a broad, fully developed measure of acceptance.

Specifically regarding the second research question, the method of assessing participants' mood is something which can be considered a limitation. Mood within this study was operationalised as mean values of participants' respective positive and negative affect scores from the PANAS over the three sessions. However, due to this only being measured at the end of each session (rather than both at the start and the end) it is likely that the measurement of mood was influenced by many factors which were not assessed by the research team. For example, measuring the influence of the LC upon participants' mood was the goal, however several other unrelated factors such as fatigue, stress level and/or appetite likely also contributed to the PANAS scores.

Avenues for Future Research

This study does not offer an ultimate conclusion on the prospect of using LCs in NFT contexts but rather it represents an important step in a feasibility framework. Future steps in this area should consider advancing the complexity of LC responses, enlarging the methodological scope of the design, and delving further into the subjective experience of the (sub-)clinical sample.

Much work is currently being undertaken to utilise artificial intelligence (AI) programmes with LC designs to enhance the range of responses they may be able to provide (Han et al., 2025). Specifically, one study has found that applications such as deep reinforcement learning (RL) can help robotic agents to interpret real-time BCI data to provide adaptive and complex feedback (Vukelic et al., 2023). It is likely that an LC which could offer more personalised and complex feedback would make individuals more likely to reach their NFT goals by improving task learning (Kochmar et al., 2020). Enhancing PU would be another likely benefit (Conati et al., 2021), something which the findings of this study highlight as an important avenue to increasing the acceptance of LCs.

Future attempts to investigate user acceptance or mood towards an LC in an NFT context should aim to extend the design of the current study in terms of both sample size and number of NFT sessions. The feasibility nature of this study meant that a very large sample size was not strictly required. However, increasing the number of participants in such a study would likely enable researchers to more thoroughly investigate how certain factors like age, (sub-)clinical status, or education might affect user attitudes towards an LC during NFT. Furthermore, although three NF sessions was considered adequate to investigate variables like acceptance and mood in this study, increasing the number of NF sessions would offer future researchers more stringent insights into variables like BI which are known to become more accurate with longer time frames (Venkatesh & Davis, 2000).

Finally, exploring the subjective experience of users in such settings is paramount to ensuring that a full-scale intervention of NFT utilising LCs will be conducted in a way which prioritises user experience. The best possible understanding of individual attitudes towards new and unfamiliar technologies can likely only be captured through both objective and subjective assessment. For example, prior literature has found focus group approaches to offer a very thorough insight into the personal and relational factors which influence BCI acceptance (Blain-Moraes et al., 2012). Implementing a subjective assessment method such as a focus group would help to ensure that the best possible insight into user attitudes towards an LC during NFT is being gained. This is of particular importance given concern that has been raised regarding the coercion of potentially vulnerable groups (such as [sub-]clinical samples) into using newer technologies against their will (Morris, 2021).

Conclusion

This study offered insight into key factors to consider when assessing the feasibility of using LCs during NFT: specifically, how interaction with them affects participants' acceptance and mood. The findings from our (sub-)clinical sample were relatively optimistic regarding acceptance but more mixed regarding mood. As is important within a feasibility context, no effects of interaction with the LC from our participants were alarming or worrying for future researchers/clinicians. Instead, the

sample was accepting of LCs in this context with some future scope for improvement of variables such as BI and PU. The interaction with the LC did not induce the mood-acceptance link as strongly as was predicted, suggesting that mood is not as important of a factor for acceptance as previous research may suggest. However, longer studies with a more rigid operationalisation of mood may be needed given the near-significant nature of the relationship which was found. Advancing the feedback abilities of LCs, extending the design and sample of the study, and incorporating subjective measures for participants all represent ways to further ensure that NFT settings are ready for LCs to become commonplace. Addressing these areas offers the opportunity to further build on the current work and to eventually maximise the effectiveness of NFT.

References

- Alexandre, B., Reynaud, E., Osiurak, F., & Navarro, J. (2018). Acceptance and acceptability criteria: a literature review. *Cognition, Technology & Work*, 20 (2), 165-177.
- Barth, B., Mayer-Carius, K., Strehl, U., Wyckoff, S. N., Haeussinger, F. B., Fallgatter, A. J., & Ehlis, A. C. (2021). A randomized-controlled neurofeedback trial in adult attention-deficit/hyperactivity disorder. *Scientific reports*, 11(1), 16873.
- Bauer, R., & Gharabaghi, A. (2015). Estimating cognitive load during self-regulation of brain activity and neurofeedback with therapeutic brain-computer interfaces. *Frontiers in behavioral neuroscience*, 9, 21.
- Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science robotics*, 3(21), eaat5954.
- Blain-Moraes, S., Schaff, R., Gruis, K. L., Huggins, J. E., & Wren, P. A. (2012). Barriers to and mediators of brain–computer interface user acceptance: focus group findings. *Ergonomics*, 55(5), 516-525.
- Brito, M. A. D., Fernandes, J. R., Esteves, N. S. A., Müller, V. T., Alexandria, D. B., Pérez, D. I. V., ... & Miarka, B. (2022). The effect of neurofeedback on the reaction time and cognitive performance of athletes: A systematic review and meta-analysis. *Frontiers in Human Neuroscience*, 16, 868450.
- Cao, J., Feng, H., Lim, Y., Kodama, K., & Zhang, S. (2024). How social influence promotes the adoption of mobile health among young adults in China: A systematic analysis of trust, health consciousness, and user experience. *Behavioral Sciences*, 14(6), 498.
- Cattell, H. E., & Mead, A. D. (2008). The Sixteen Personality Factor. *The SAGE Handbook of Personality Theory and Assessment: Personality Measurement and Testing (Volume 2)*, 2, 135.
- Chao, C. M. (2019). Factors determining the behavioral intention to use mobile learning: An application and extension of the UTAUT model. *Frontiers in psychology*, 10, 1652.

- Chou, C. Y., Chan, T. W., & Lin, C. J. (2003). Redefining the learning companion: the past, present, and future of educational agents. *Computers & Education*, 40(3), 255-269.
- Conati, C., Barral, O., Putnam, V., & Rieger, L. (2021). Toward personalized XAI: A case study in intelligent tutoring systems. *Artificial intelligence*, 298, 103503.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340.
- De Graaf, M. M., & Allouch, S. B. (2013). Exploring influencing variables for the acceptance of social robots. *Robotics and autonomous systems*, 61(12), 1476-1486.
- Earle-Randell, T. V., Wiggins, J. B., Ma, Y., Celepkolu, M., Bounajim, D., Gao, Z., ... & Wiebe, E. (2024). The impact of near-peer virtual agents on computer science attitudes and collaborative dialogue. *International Journal of Child-Computer Interaction*, 40, 100646.
- Edwards, A., Edwards, C., Abendschein, B., Espinosa, J., Scherger, J., & Vander Meer, P. (2022). Using robot animal companions in the academic library to mitigate student stress. *Library Hi Tech*, 40(4), 878-893.
- Eisenbarth, H., D'Cruz, C., Bulbulia, J. A., & Thanni, B. (2025). Culturally Diverse Perceptions of EEG and Neurofeedback Research and How to Address Them to Reduce Sampling Bias. *Psychophysiology*, 62(6), e70077.
- Fernández-Álvarez, J., Grassi, M., Colombo, D., Botella, C., Cipresso, P., Perna, G., & Riva, G. (2022). Efficacy of bio-and neurofeedback for depression: a meta-analysis. *Psychological medicine*, 52(2), 201-216.
- Fredericks, S., Sidani, S., Fox, M., & Miranda, J. (2019). Strategies for balancing internal and external validity in evaluations of interventions. *Nurse Researcher*, 27(4).
- Gadke, D. L., Kratochwill, T. R., & Gettinger, M. (2021). Incorporating feasibility protocols in intervention research. *Journal of School Psychology*, 84, 1-18.

- Grevet, E., Forge, K., Tadiello, S., Izac, M., Amadiou, F., Brunel, L., ... & Jeunet-Kelway, C. (2023). Modeling the acceptability of BCIs for motor rehabilitation after stroke: A large scale study on the general public. *Frontiers in Neuroergonomics*, 3, 1082901.
- Grevet, E., Izac, M., Amadiou, F., Py, J., Gasq, D., & Jeunet-Kelway, C. (2024, May). Which factors affect patients' acceptability of BCIs for functional rehabilitation after stroke? A questionnaire study among 140 patients and a comparison with the general public. In *Journées CORTICO Scientific Days 2024*.
- Han, Y., Hong, S., Li, Z., & Lim, C. (2025). Defining and Classifying the Roles of Intelligent Learning Companion Systems: A Scoping Review of the Literature. *TechTrends*, 1-15.
- Hasslinger, J., D'Agostini Souto, M., Folkesson Hellstadius, L., & Bölte, S. (2020). Neurofeedback in ADHD: A qualitative study of strategy use in slow cortical potential training. *PLoS One*, 15(6), e0233343.
- Heubeck, B. G., & Wilkinson, R. (2019). Is all fit that glitters gold? Comparisons of two, three and bi-factor models for Watson, Clark & Tellegen's 20-item state and trait PANAS. *Personality and Individual Differences*, 144, 132-140.
- Jackson, L. E., Han, Y. J., & Evans, L. H. (2023). The efficacy of electroencephalography neurofeedback for enhancing episodic memory in healthy and clinical participants: A systematic qualitative review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, 155, 105455.
- Jeunet, C., N'Kaoua, B., & Lotte, F. (2016). Advances in user-training for mental-imagery-based BCI control: Psychological and cognitive factors and their neural correlates. *Progress in brain research*, 228, 3-35.
- Johnson, W. L., & Lester, J. C. (2018). Pedagogical agents: back to the future. *AI Magazine*, 39(2), 33-44.
- Kadosh, K. C., & Staunton, G. (2019). A systematic review of the psychological factors that influence neurofeedback learning outcomes. *Neuroimage*, 185, 545-555.
- Kapoor, A., Mota, S., & Picard, R. W. (2001, November). Towards a learning companion that recognizes affect. In *AAAI Fall symposium* (Vol. 543, pp. 2-4).

- Kim, Y., & Baylor, A. L. (2006). A social-cognitive framework for pedagogical agents as learning companions. *Educational technology research and development*, 54(6), 569-596.
- Kochmar, E., Vu, D. D., Belfer, R., Gupta, V., Serban, I. V., & Pineau, J. (2020, June). Automated personalized feedback improves learning gains in an intelligent tutoring system. In *International conference on artificial intelligence in education* (pp. 140-146). Cham: Springer International Publishing.
- Kort, B., Reilly, R., & Picard, R. W. (2001, August). An affective model of interplay between emotions and learning: Reengineering educational pedagogy-building a learning companion. In *Proceedings IEEE international conference on advanced learning technologies* (pp. 43-46). IEEE.
- Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of experimental social psychology*, 49(4), 764-766.
- Lester, J. C., Towns, S. G., & Fitzgerald, P. J. (1998). Achieving affective impact: Visual emotive communication in lifelike pedagogical agents. *International Journal of Artificial Intelligence in Education*, 10, 278-291.
- Liew, T. W., Mat Zin, N. A., & Sahari, N. (2017). Exploring the affective, motivational and cognitive effects of pedagogical agent enthusiasm in a multimedia learning environment. *Human-centric Computing and Information Sciences*, 7(1), 9.
- Lim, C. G., Soh, C. P., Lim, S. S. Y., Fung, D. S. S., Guan, C., & Lee, T. S. (2023). Home-based brain-computer interface attention training program for attention deficit hyperactivity disorder: A feasibility trial. *Child and Adolescent Psychiatry and Mental Health*, 17(1), 15.
- Loriette, C., Ziane, C., & Hamed, S. B. (2021). Neurofeedback for cognitive enhancement and intervention and brain plasticity. *Revue Neurologique*, 177(9), 1133-1144.
- Lotte, F., Larrue, F., & Mühl, C. (2013). Flaws in current human training protocols for spontaneous brain-computer interfaces: lessons learned from instructional design. *Frontiers in human neuroscience*, 7, 568.

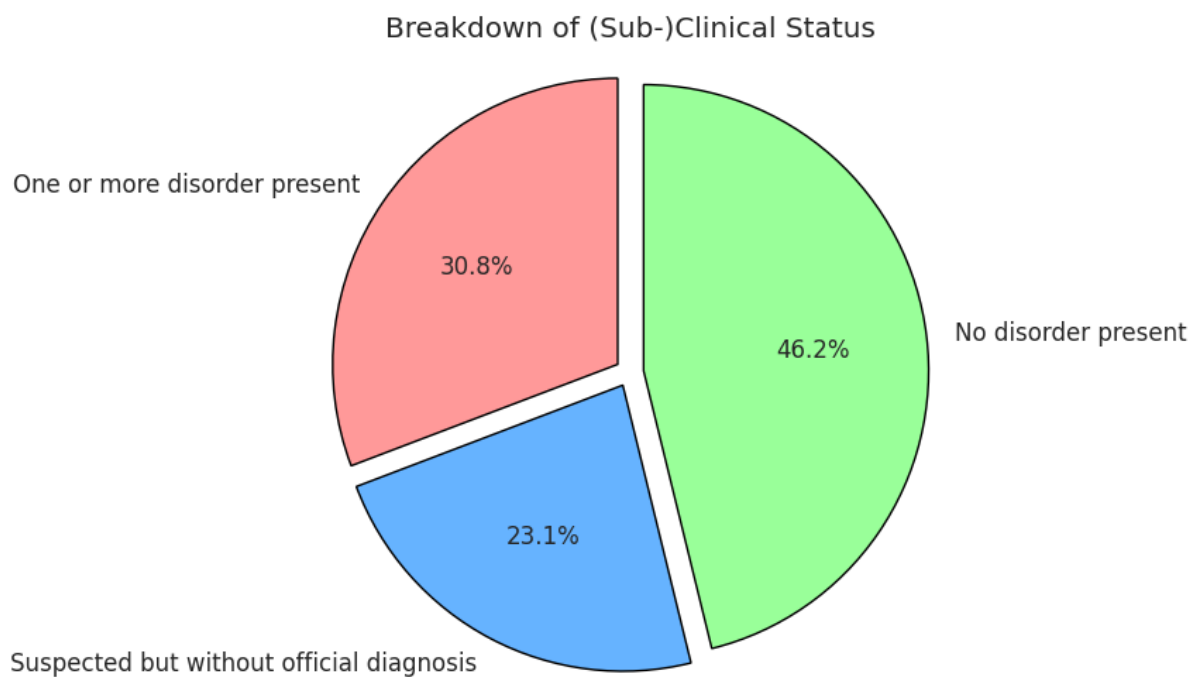
- Mansour, S., Giles, J., Nair, K. P., Marshall, R., Ali, A., & Arvaneh, M. (2025). A clinical trial evaluating feasibility and acceptability of a brain-computer interface for telerehabilitation in patients with stroke. *Journal of NeuroEngineering and Rehabilitation*, 22(1), 91.
- Morris, N. P. (2021). Digital technologies and coercion in psychiatry. *Psychiatric services*, 72(3), 302-310.
- National Institute for Health Research. (2012). NIHR Evaluation, Trials and Studies Coordination Centre: Glossary.
- Nijboer, F., Birbaumer, N., & Kübler, A. (2010). The influence of psychological state and motivation on brain-computer interface performance in patients with amyotrophic lateral sclerosis—a longitudinal study. *Frontiers in Neuropharmacology*, 4, 55.
- Ortiz-Ordoñez, J. C., Stoller, F., & Remmele, B. (2015). Promoting self-confidence, motivation and sustainable learning skills in basic education. *Procedia-Social and Behavioral Sciences*, 171, 982-986.
- Pillette, L., Jeunet, C., N'Kambou, R., N'Kaoua, B., & Lotte, F. (2019). Towards artificial learning companions for mental imagery-based brain-computer interfaces. *arXiv preprint arXiv:1905.09658*.
- Pillette, L., Jeunet, C., Mansencal, B., N'kambou, R., N'Kaoua, B., & Lotte, F. (2020). A physical learning companion for Mental-Imagery BCI User Training. *International Journal of Human-Computer Studies*, 136, 102380.
- Rosli, M. S., Saleh, N. S., Md. Ali, A., Abu Bakar, S., & Mohd Tahir, L. (2022). A systematic review of the technology acceptance model for the sustainability of higher education during the COVID-19 pandemic and identified research gaps. *Sustainability*, 14(18), 11389.
- Roth, R. M., Isquith, P. K., & Gioia, G. A. (2005). Behavior rating inventory of executive function®--adult version. *Archives of clinical neuropsychology*.
- Roth, R. M., Lance, C. E., Isquith, P. K., Fischer, A. S., & Giancola, P. R. (2013). Confirmatory factor analysis of the behavior rating inventory of executive function-adult version in healthy adults and application to attention-deficit/hyperactivity disorder. *Archives of clinical neuropsychology*, 28(5), 425-434.

- Schreibelmayr, S., & Mara, M. (2022). Robot voices in daily life: Vocal human-likeness and application context as determinants of user acceptance. *Frontiers in psychology*, 13, 787499.
- Share, P., & Pender, J. (2018). Preparing for a robot future? Social professions, social robotics and the challenges ahead. *Irish Journal of Applied Social Studies*, 18(1), 4.
- Song, Y., Luximon, A., & Luximon, Y. (2021). The effect of facial features on facial anthropomorphic trustworthiness in social robots. *Applied Ergonomics*, 94, 103420.
- Stommel, WJ, & Stommel, MW (2021). Participation of companions in video-mediated medical consultations: A microanalysis. In *Analyzing digital interaction* (pp. 177-203). Cham: Springer International Publishing.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International journal of medical education*, 2, 53.
- Thompson, E. R. (2007). Development and validation of an internationally reliable short-form of the positive and negative affect schedule (PANAS). *Journal of cross-cultural psychology*, 38(2), 227-242.
- Tolin, D. F., Davies, C. D., Moskow, D. M., & Hofmann, S. G. (2020). Biofeedback and neurofeedback for anxiety disorders: a quantitative and qualitative systematic review. *Anxiety disorders: Rethinking and understanding recent discoveries*, 265-289.
- TULIP Study. (n.d.). NEUROFEEDBACK. Retrieved April 16, 2025, from <https://www.neurofeedback-rug.nl/tulip-study.html>
- Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management science*, 46(2), 186-204.
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425-478.

- Vlachou, J. A., Polychroni, F., Drigas, A. S., & Economou, A. (2022). Neurofeedback and ADHD. *International Journal of Recent Contributions from Engineering, Science & IT (iJES)*, 10(01), 47-56.
- Vukelić, M., Bui, M., Vorreuther, A., & Lingelbach, K. (2023). Combining brain-computer interfaces with deep reinforcement learning for robot training: a feasibility study in a simulation environment. *Frontiers in neuroergonomics*, 4, 1274730.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology*, 54(6), 1063.
- Zielke, M. A., Zakhidov, D., Lo, T., Craig, S. D., Rege, R., Pyle, H., ... & Kuo, N. (2025). Exploring social learning in collaborative augmented reality with pedagogical agents as learning companions. *International Journal of Human–Computer Interaction*, 41(4), 2424-2449.

Appendices

Appendix A1: Breakdown of (sub-)Clinical Status

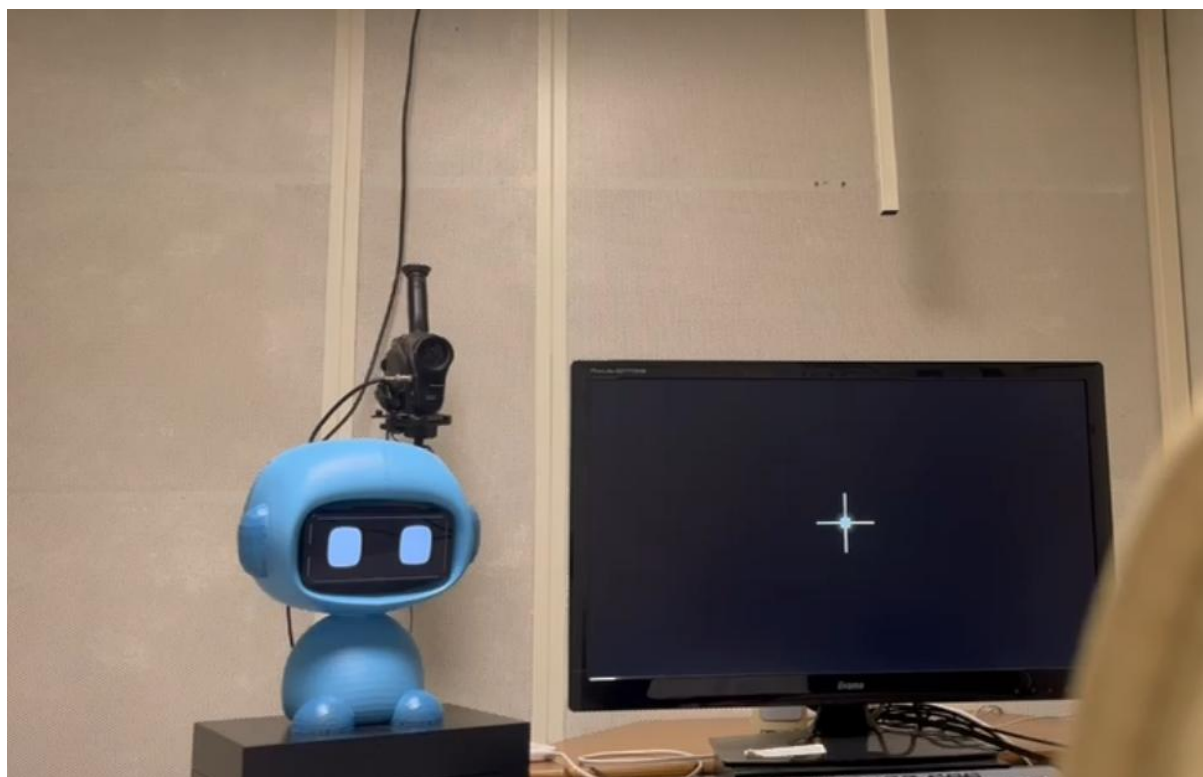


Appendix A2: Breakdown of (sub-)Clinical Status

Participant ID	Pre-assessment questionnaire questions		
	Do you have one or more psychiatric or neurological disorders?	If so, what are they?	Which cognitive challenge affects your daily life the most?
A2309C	No		Sustained attention
P2511W	Yes	ADHD, ASD	Flexibility, prioritisation and attention
T0709D	Yes	ADHD, GAD	Task initiation, time management, difficulty planning and prioritising, maintaining attention
M2904H	I suspect, not officially diagnosed	ADHD, GAD	Emotion regulation, working memory, response inhibition, easily overstimulated, fluctuating energy
O2502S	No		Attention, lack of interest
I1308S	No		Task initiation
k2408g	No		Planning and prioritising
S2501N	Yes	ASD, OCD	Emotion regulation, task initiation, flexibility
E2706K	I suspect, not officially diagnosed		Attention, difficulty planning/prioritising, working memory
K0304A	No		Time management
t2010s	Yes	MS	Planning, time management, organisation
V1411H	No		Time management, difficulties planning/prioritising
S0108K	I suspect, not officially diagnosed		Planning/prioritising, emotion regulation, time management

Appendix B

Appendix B1: Image of the LC (Juno) sitting left of the NF screen



Appendix C

Appendix C1: Mental Strategies List

Neurofeedback

The goal of neurofeedback training is to alter your brain activity using a range of mental strategies. The actual mechanisms behind why the brain activity is altered is unknown and the strategies employed to trigger this change vary from participant to participant. Within this study, you can monitor the efficacy and degree of the desired changes by employing a variety of different mental strategies, some of which are listed below. The effect of each mental strategy will be represented with a square on the screen during the training. The square will turn green if a strategy is working and keeping this square green for a long time is a good indication that a strategy is effective. So, attempt some of the strategies listed below or try any strategy that you can come up with; the most important thing is that you find one that works well for you.

Possible Neurofeedback Strategies

Mental tasks, such as:

- Perform arithmetic/calculation tasks
- Recall memory content (e.g. what did I eat this week or what people did I meet last month?)
- Planning (e.g. what will I do next week?)
- Mentally rotate objects (e.g. imagine rotating a cup 360 degrees)
- Mentally navigate (through familiar buildings or streets)
- Allowing spatial attention to wander (e.g. moving to different areas without losing sight of the space (from one side to the other))

Relaxation, e.g. through concentration on breathing

Imagining emotions (positive/negative)

Recall memories/imagining situations

- Family members (parents, siblings, grandparents)
- Friends and acquaintances
- Partners

Auditory imaginations, e.g. imagining sounds or music

Cheering on the green square

Imagining movement or activities

- Arm or foot movements
- Playing sports
- Singing
- Movie watching

Thoughts about nature/imagination of nature

- Imagine it raining
- Imagine a sunset
- Imagine certain landscapes
- Imagine a journey

Thoughts/imagining everyday things

- Cooking or eating
- Going shopping
- Cleaning

Most important: [try your own strategies](#)

Appendix D

Appendix D1: Pre and Post intercorrelations of BI, PEOU and PU

Variable	BI	PEOU	PU
Pre			
BI	-	.117	.644
PEOU	.117	-	.3
PU	.644	.3	-
Post			
BI	-	.44	.79
PEOU	.44	-	.34
PU	.79	.34	-

Appendix D2: Pre and Post Internal Consistency (Cronbach's α)

Variable	No. of Items	α pre	α post
BI	3	0.821	0.926
PEOU	3	0.473	0.216
PU	3	0.899	0.891
Acceptability/Acceptance	9	0.782	0.874

Appendix E

Appendix E1: Breakdown of results from paired samples *t*-tests for SC, FC, SI and ID

Variable	<i>T</i> (df)	<i>p</i>	<i>p</i> [*]	<i>Cohen's d</i>
SC	0.6(12)	0.28	1	0.14
FC	0.61(12)	0.554	1	0.1
SI	-0.18(12)	0.864	1	0.04
ID	-4.05(12)	0.999	1	1.16

Note. *=value corrected via Holm-Bonferroni correction

Appendix E2: Descriptive breakdown of means and standard deviations for SC, FC, SI, and ID.

Variable	<i>M</i>		<i>SD</i>		
	n	Pre (session 1)	Post (session 3)	Pre (session 1)	Post (session 3)
SC	13	49.35	51.24	13.06	13.46
FC	13	53.55	54.66	10.49	10.82
SI	13	64.10	63.51	16.55	10.8
ID	13	51.82	41.23	9.79	8.38

Appendix F:

AI system: ChatGPT, 2025

Modifications: No modifications were used but AI was only used for coding/statistical analysis purposes.

Use case: used for generating outputs of code in Python to help with statistical analysis.

Final prompts used: *“how do I group the outputs of multiple subsets into one new data set?”, “how do I add a legend which describes each of the sets of data in my plot”, and “when plotting a line plot in APA 7 where the y-axis starts above 0 (it starts at 30), do I need to add ticks on every score on that axis?”*