Assessing Practitioners' Acceptability of a Robot-Like Neurofeedback Companion: Implications for Clinical Implementation

Josephine Mathing

S5049385

Department of Psychology, University of Groningen

PSB3E-BT15: Bachelor Thesis

Group number: 31

Supervisor: Dr. Stefanie Enriquez-Geppert

Second evaluator: Prof. Dr. Ralf Cox

In collaboration with: Alberta-Maria Lupașcu, Esmee van der Veen, Ilse Dusseltje, Marrit Wierda, Sara Maria Juganaru

July 1, 2025

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.

I did not use AI to produce any content in this thesis. AI (Grammarly) was used for grammar checking throughout the entire thesis. Additionally, ChatGPT (OpenAI, 2023) was used for synonym generation and occasional translation from German to English. Example prompts used are: “What can I use instead of additionally or moreover?”, “Translate: Ein Ansatzpunkt”.

The maximum word count of this thesis was extended in accordance with our supervisor.

Abstract

Introduction: Neurofeedback (NF) training is an intervention developed to help people regulate their brain activity. However, not everyone responds to this training. Thus, researchers are looking to decrease the number of non-responders by adding a robot-like NF companion to enhance the traditional feedback provided. Therefore, this feasibility study assesses practitioners' acceptance of a new robot-like NF companion. By assessing practitioners' acceptance specifically, we aim to enhance the future implementation of the robot-like NF companion since practitioners are the ones to use it in practice. **Methods:** A convenience sample of $n = 618$ participants, including practitioners with and without prior knowledge of NF and the general public, was utilized in this analysis. All participants filled out an online questionnaire assessing different aspects of acceptability concerning the robot-like NF companion. The Behavioral Intention (BI) component was of interest in this paper. The analyses included an independent samples t-test, a Mann-Whitney U test, and a three-way between-subjects ANOVA. **Results:** Generally, high levels of BI scores were found across groups. Practitioners scored lower in BI than the general public in a matched sample condition. There was no significant difference between age, gender, and experience levels. **Discussion:** The results indicate a favorable attitude towards the robot-like NF companion; therefore, further testing of it in actual NF sessions should be conducted. Practitioners should be involved in the next steps of testing and in focus groups to enhance future implementation and allow space for the expression of concerns in order to bring the robot-like NF companion into clinical practice.

Keywords: Neurofeedback, Acceptability, Feasibility Study, Learning Companion, Research-to-Practice Gap, Implementation Science, Practitioners

Assessing Practitioners' Acceptability of a Robot-Like Neurofeedback Companion: Implications for Clinical Implementation

Neurofeedback (NF) training is a promising technique for helping people self-regulate their brain activity based on operant conditioning (Viviani & Vallesi, 2021). It can be applied in various contexts, including people with cognitive difficulties (Enriquez-Geppert et al., 2014), various psychiatric disorders (Pindi et al., 2022; Van Doren et al., 2018), and otherwise healthy people (Matsuzaki et al., 2023). NF training can be conducted in various settings like specialized clinics and practices, but should always be conducted by a licensed professional with NF experience (Hammond, 2007). During the non-invasive training, the brain waves are measured using, for example, electroencephalography (EEG). Based on the measured brain waves, feedback is fed back to the participant in the form of real-time video or audio feedback, allowing the participant to regulate their brain activity (Marzbani et al., 2016; Viviani & Vallesi, 2021). This regulation in brain activity can lead to a reduction of cognitive symptoms (Smit et al., 2023). Often-used visual feedback during NF training includes a moving bar against a black background with changing colors on a screen in front of the participant (Mathiak et al., 2015). Nevertheless, there are some people who do not respond to NF training.

These people are called non-responders and comprise 15-30% of those who undergo NF training (Pillette et al., 2020). The non-responders are not able to successfully control their brainwaves with NF and therefore do not benefit from it (Pillette et al., 2020). It is possible that these people simply do not respond to the traditional visual feedback that is provided. Shute (2008) pointed out that feedback should be, amongst other things, non-evaluative and supportive. Thus, adjusting and expanding on the traditional feedback could be a way to reduce the number of non-responders.

One idea that could improve the feedback is the use of a robot-like NF companion. This type of robot-like NF companion is based on learning companions, which can be defined as “a (computer-simulated) character, which has human-like characteristics and plays a non-authoritative role in a social learning environment” (Chou et al., 2003, p.258). It aids students during various learning activities. An example of one such learning companion is “Minnie”, who was developed by Michaelis and Mutlu (2017). Minnie was designed to be an in-home learning companion robot for children, helping to foster engagement with reading. Minnie provided individualized verbal and non-verbal feedback that was based on the child's preferences as well as serving as a social partner during reading (Michaelis & Mutlu, 2017). Building on this, Pillette et al. (2020) developed a robot-like learning companion called “PEANUT” to implement during Mental-Imagery based Brain-computer Interface training (MI-BCI). MI-BCI is similar to NF since users use their brain activity to send commands to digital technologies (Pillette et al., 2020). PEANUT was developed to improve the type of feedback during the MC-BCI sessions and provide interventions based on the participants' performance and progression. PEANUT added to the standard visual feedback provided during MI-BCI by providing verbal and non-verbal feedback based on the participant's data from both current and previous sessions. This approach yielded beneficial results in healthy participants regarding performance and user experience (Pillette et al., 2020). Additionally, one study found that implementing social reward (e.g., an avatar smiling when brain activity increases) during NF training led to stronger activation of the brain area of interest during training and showed a trend for better performance after the session when no social reward as feedback was provided (Mathiak et al., 2015). Social rewards stimulate the same brain regions associated with rewards as other kinds of rewards (Izuma et al., 2008) and have been proven to directly strengthen localized brain activity (Mathiak et al., 2010). Based on these findings, developing a robot-like NF companion with social rewards and individualized feedback based on progress and performance could be beneficial for NF training. However, one big hurdle for the implementation is the so-called “research-to-practice gap”.

This gap illustrates the differences between what is researched, published in journals, and accepted by most researchers, and what is helpful when it comes to practice (Keegan et al., 2017). Sometimes this gap is accidental, due to miscommunication, other times it is deliberate, due to incorrect application or research being irrelevant to practice (Norman, 2010). As some researchers put it, “researchers ‘get on with’ research, and the practitioners ‘get on with’ practice, solving different problems using different methods and approaches; separately” (Keegan et al., 2017, pp. 75-76). It was estimated by Balas and Boren (2000) that it takes 17 years for 14% of research to be implemented in clinical practice. This problem is especially visible in clinical neuropsychology. For example, a recent study found that only 20% of memory clinics in Australia provided cognitive interventions to patients, even though 74% of respondents identified these types of interventions as important (Naismith et al., 2022). Born out of the need to solve this problematic gap, a new branch of science was created. It is called Implementation Science and investigates the most promising ways to integrate evidence-based findings into clinical practice (Olswang & Prelock, 2015). Currently, more than 600 determinants that could be important in the implementation of various evidence-based findings have been identified (Williams & Beidas, 2018). Regarding the implementation of a robot-like NF companion, several factors could influence a practitioner's willingness to implement them.

The gender of the practitioner could play a role in this decision. Venkatesh et al. (2000) found that gender influences the initial decision-making process that determines the adoption of new technology, with men having a more salient attitude towards using technology. Moreover, in a study examining higher education students’ attitudes towards using technology, it was found that male students were more confident in using technology than female students (Yau & Cheng, 2012). These findings indicate that male practitioners could be more willing to implement and work with the robot-like NF companion.

Apart from gender, age could be a significant next factor. Older adults have been found to frequently adopt technology less than younger generations (Anderson & Perrin, 2017). Some factors that play a role in the decreased willingness to adopt technology are a lack of awareness and experience, decreased confidence, low perceived value, and negatively self-assessed abilities to use technology (Berkowsky et al., 2018). This indicates that younger practitioners may be more willing to implement the new robot-like NF companion in their practice.

Overall, according to the Unified Theory of Acceptance and Use of Technology 2, age and gender play important moderating roles in the acceptance of new technologies (Venkatesh et al., 2012) and therefore underline the importance of considering those variables in the implementation of the robot-like NF companion. Moreover, while not being explicitly part of the model, age and gender differences have been found for the Technology Acceptance Model (Assaker, 2019), which was developed to understand the acceptance of IT technologies (Holden & Karsh, 2010).

Specifically, regarding therapists' attitudes towards new technologies, Liu et al. (2014) found that the current use of new technologies could be predicted accurately from the behavioral intention to use a new technology. Moreover, the higher the belief that the new technology will help to improve job performance or patient outcomes, the higher the intention to use it (Liu et al., 2014). Additionally, research indicates that therapists who are more open to using new treatments also have stronger intentions to use new technologies (Ahuna et al., 2023). Similarly, practitioners with more clinical experience are more inclined to use evidence-based therapies in practice (Beidas et al., 2017). All in all, technology openness, the perceived benefit of the new technology, clinical experience, and the behavioral intention to use a new technology seem to be important factors when evaluating the possible implementation of a robot-like NF companion.

To help practitioners make decisions about the implementation of new technologies, feasibility studies can be employed. Bowen et al. (2009) developed a framework for feasibility studies. According to them, a feasibility study is “used to determine whether an intervention is appropriate for further testing” (Bowen et al., 2009, p.2). These types of studies can help develop new techniques or therapies and aid in their possible future implementation by guiding the whole process and enabling the researchers to implement changes in their methods or protocols along the way (Bowen et al., 2009). There are eight areas a feasibility study can focus on. The first one, and the one of importance for this paper, is acceptability.

Acceptability, in a feasibility study, examines how individuals who are intended to use the intervention and individuals who implement and administer the intervention react to the proposed intervention (Bowen et al., 2009). Acceptability has several components and deals with the view of an individual before any interaction with the intervention has occurred (Grevet et al., 2023). Grevet et al. (2023) developed an acceptability model for BCIs based on several models, including the Technology Acceptance Model 3 and the Unified Theory of Acceptance and Use of Technology 2. This model included behavioral intention (BI), perceived usefulness, and perceived ease of use as components of acceptability, which are also components of the other two models. The BI component of acceptability is of special importance to practitioners, as they are the ones who ultimately can decide to use an intervention, like a robot-like NF companion during NF training, in practice. Moreover, BI has been found to reliably predict the actual future use (Holden & Karsh, 2010) and is therefore a crucial component that should be adequately assessed when investigating the possible implementation of new technologies.

To assess practitioners' willingness to implement the robot-like NF companion, this paper investigates whether there is a difference in the BI to use a robot-like NF companion for neurofeedback training between practitioners who are already working with neurofeedback and those who are not, compared to the general public. Practitioners who already work with NF are defined as people with a professional background in the medical or health-related fields who include/included NF in their therapies. Practitioners with no prior experience with NF include neurologists, psychiatrists, and nurses. The "general public" was an age- and gender-matched group to the practitioners with no background in healthcare. These hypotheses were formulated based on the research question above:

H1: Practitioners in general will have a higher intention to use a robot-like NF companion than the general public.

$H0_1$: There will be no difference in the intention to use a robot-like NF companion between practitioners and the general public.

H2: Practitioners already working with neurofeedback will have a higher intention to use a robot-like NF companion than practitioners who do not.

$H0_2$: There will be no difference in the intention to use a robot-like NF companion between practitioners working with neurofeedback and practitioners who do not.

H3: Younger practitioners will have a higher intention to use a robot-like NF companion than older practitioners. This relationship will be stronger among those with NF experience compared to those without, which will be even more pronounced by gender.

$H0_3$: There will be no difference in the intention to use a robot-like NF companion between ages, NF experience, and gender.

Methods

Recruitment and inclusion criteria

The recruitment of the participants has been done through the network of bachelor students and the TULIP research team, the use of flyers and social media, through writing emails to professionals who work in a psychological context, and the SONA system. The minimum age for participating was eighteen years old, and understanding one of the following languages was crucial: Dutch, English, Spanish, German, or French, as the questionnaire was available in these languages. Participation was voluntary and unpaid. The collection of the data started in February 2024 and is still running. The study questionnaire (PSY-2324-S-0092 TULIP-acceptability study) was approved by the Ethical Committee of the Behavioral and Social Science Faculty of the University of Groningen, Netherlands, and conducted in accordance with the Declaration of Helsinki.

Participants

This study used a convenience sample. There were 854 adult participants collected from different countries. The mean age was 29.03 years (*SD* = 15.04). The gender distribution is: 65.81% of the participants were women (*n* = 562), 20.26% were men (*n* = 173), 2.93% fell in the category "other" (*n* = 25), and there were 11.01% of people who did not fill in anything regarding gender (*n* = 94). Participants were provided with information about their involvement before taking part in the study, as well as a description of how their data would be used following the ethical principles that protect the rights and well-being of participants. They gave informed consent to participate.

Procedure and Materials

The questionnaire was created and distributed using Qualtrics. The participants completed the questionnaire on their own devices. The estimated duration required to complete the questionnaire is approximately thirty minutes. They had to sign the informed consent form in order to start the questionnaire. The questionnaire contained questions about demographics, cognitive problems, NF understanding and attitudes, the robot-like NF companion's trustworthiness and acceptability, and personality traits.

Demographics Questionnaire

In the first part of the survey, questions related to sociodemographic background (age, gender, work status, completed education, residency, nationality, medical or health-related profession, history of psychiatric or neurological conditions) were administered.

Questions Regarding Experienced Cognitive Problems

Cognitive problems were measured using 14 items in which participants rated their response on a 3-point scale: *Yes, strongly*, *Yes, slightly*, or *No*. Multiple facets of cognitive problems were assessed, with each item introduced by a specific cognitive domain label, followed by a descriptive statement. The last item provided participants with the opportunity to report cognitive concerns not previously covered. An additional question was addressed regarding engagement in cognitive abilities training, in which participants rated their response on a scale from 1 to 100, with 1 indicating *Not at all* and 100 indicating *Yes a lot*.

Neurofeedback Educational Information

Participants received some information on neurofeedback. This also included some images that further explained this process. Participants also received some basic information on the goals of the neurofeedback companion. They were also shown a short video of the setup, including the learning companion. After reviewing all the information, participants had to complete a short quiz to ensure understanding and continue with the study. After the quiz, participants had to answer a few questions on whether they felt like they understood what neurofeedback and the learning companion are.

Experiences, Attitudes, and Expectations Regarding Neurofeedback and Related Techniques

Participants were asked about whether they had heard about neurofeedback before participating in this study, and if they had any previous experience with it. Finally, they were asked questions on a 0-100 scale on their attitudes and expectations regarding NF, with 0 being *completely disagree*, and 100 being *completely agree*.

Questionnaire Assessing the Design Preferences of the Neurofeedback Learning Companion

Participants were asked to indicate whether they had ever been in contact with a learning companion before, and if so, what kind of companion it was.

They were asked to rank four different companions based on the trustworthiness of their shape based on pictures included in the questionnaire. Moreover, participants were asked to indicate how trustworthy they found four possible names of the companion. Trustworthiness was indicated on a scale from 0 (*not trustworthy*) to 100 (*maximally trustworthy*). With trustworthiness, we mean that participants would accept the learning companion's feedback and apply it during the neurofeedback sessions.

The type of color and the number of colors a trustworthy companion should have were evaluated as well. Participants were able to choose between four different options each. Additionally, participants were asked to assess four voice samples and indicate which voice a trustworthy companion should have. The perceived fit of the voice was indicated on a scale from 0 (*not fitting at all*) to 100 (*maximally fitting*).

The behavior of the companion was evaluated as well. Participants were asked when they would like to receive feedback from the companion during the neurofeedback sessions (e.g., only when they succeed, only when they fail, or both) and whether they thought that a companion could distract them from neurofeedback. This last point was assessed using a scale from 0 (*completely disagree*) to 100 (*completely agree*).

Acceptability of Neurofeedback Companion Questionnaire

Acceptability was measured using eleven selected and adapted questions from the Acceptability Model for BCIs designed by Grevet et al. (2023). Within this variable, three key components were assessed using three questions for each: perceived ease of use, perceived usefulness, and BI.

In addition to the key components of acceptability, the questionnaire also included one item each for both technology-related pleasure and confidence in using new technologies, suggesting the importance of psychological factors in acceptability (Grevet et al., 2023). All eleven selected questions were in the form of a visual analogue scale ranging from 0 (*totally disagree*) to 100 (*totally agree*). Based on this, a score of 60 was set by us as the threshold to assess a positive BI. This means that a participant had a favorable attitude towards the companion and would possibly be willing to use it in the future.

Grevet et al. 's (2023) Acceptability Model has great internal consistency as demonstrated by the Cronbach's α scores ranging from .83 to .97 for the subdomains of acceptability. Regarding attitudes on technology, while the domain of pleasure shows a high level of internal consistency ($\alpha = .83$), confidence in using new technologies is rated more poorly ($\alpha = .57$).

Big Five Inventory

Personality was assessed using the Big Five Inventory (John et al., 1991), which consists of 44 items rated on a 5-point Likert scale ranging from *Disagree strongly* to *Agree strongly*. The questionnaire measures five personality domains, namely openness to experience, conscientiousness, agreeableness, extraversion, and neuroticism. The instrument demonstrates a good internal consistency, with Cronbach's α values ranging from .79 (Agreeableness) to .88 (Extraversion), and an average α of .83 across all domains (John & Srivastava, 1999). The personality scores were created by adding up the scores of the questions for each trait.

Statistical Analysis

The first hypothesis expected that practitioners would have a higher BI to use a robot-like NF companion than the general public. An independent samples t-test was conducted.

The second hypothesis proposed that practitioners with prior NF experience would have a higher BI to use the robot-like NF companion than practitioners without prior NF experience. An independent samples t-test was conducted.

The third hypothesis assumed that younger practitioners would have a higher BI than older practitioners, with the relationship being stronger for those with prior NF experience compared to those without, and with the relationship being more pronounced by gender. Thus, a three-way between-subjects Analysis of Variance (ANOVA) was conducted.

All assumptions for both the t-tests and the ANOVA were tested prior to the analyses. Normality was assessed using the Shapiro-Wilk test and examining Q-Q plots. Homogeneity of Variances was assessed using Levene's test. For all inferential tests, a significance level of $\alpha = .05$ was used. Effect sizes were reported for every analysis to bolster significance testing. All statistical analyses were conducted using IBM SPSS Statistics v.28.

Data preparation

The original data set included 854 cases. From those cases, the ones who did not fill in any age ($n = 2$), the ones who fit into the gender category "other" ($n = 16$), and the ones who did not fill out the acceptability part of the questionnaire ($n = 218$) were removed from the data set for this analysis. After the removal, the data set contained 618 cases.

A new variable called *BI_average* was calculated from the mean score of the three questions of this construct in the questionnaire (*BI_1, BI2_1, BI3_1*). This variable was used as the dependent variable in all analyses.

Hypothesis 1. To be able to conduct the independent samples t-test, a new variable called *healthcare_bg* was created. The variable *Q6.1* (which indicates whether a participant has a background in healthcare) was recoded into the new one with 1 = practitioner (n = 101) and 2 = general public (n = 517).

To accurately compare the two groups, it was decided to match the practitioner group by age and gender to the general public group. To do so, the Case Control Matching function of SPSS was used. The variable that was matched on was the *BI_average* variable, and the variable *healthcare_bg* was used as the grouping indicator. The match tolerances for both age and gender were set as 0. This matching resulted in 62 exact matches by age and gender, with 39 cases remaining unmatched. It was decided to exclude the 39 unmatched cases and not use fuzzy matches to achieve the most accurately matched groups for the analysis.

Hypothesis 2. To be able to conduct the independent samples t-test, a new numeric variable called *NF_experience* (1 = experience, 2 = no experience) was created from the string variable *Q13.4,* which indicated whether practitioners already had prior NF experience (included NF in their therapies). This question (*Q13.4*) was only shown to participants with a background in healthcare and who had already heard of NF prior to this questionnaire. The final sample for testing this hypothesis included 52 practitioners with no prior NF experience and 22 practitioners with NF experience (combined n = 74).

Hypothesis 3. The three-way between-subjects ANOVA examined the main effects of age (younger/older), NF experience (yes/no), and gender (male/female). Moreover, three two-way interactions were examined: age x NF experience, age x gender, and NF experience x gender. Lastly, the three-way interaction of age x NF experience x gender was analyzed.

To do so, a new numeric variable called *gender* (1 = female, 2 = male) was created. For this analysis, the female group included 60 cases, while the male group included 14 cases. Additionally, a numeric variable called *age_group* (1 = younger, 2 = older) was created based on the median age of the practitioners (*Mdn* = 42yrs). The younger group (<= 42yrs) encompassed 30, and the older group (> 42yrs) 44 cases. The variable *NF_experience*, with the same number of cases as in the second hypothesis, was also utilized in this analysis.

Results

Sample Characteristics

The characteristics of the final sample (*n* = 618) are listed below in Table 1. The mean age of the final sample was (*M* = 27.6, *SD* = 14.5), with 76.7% of the sample being women and 16.3% being practitioners. The mean average of BI for the final sample was (*M* = 67.1, *SD* = 21.0). Moreover, an overview of the BI average per variable is provided in Table 2.

Table 1

Sample Characteristics After Dataset Reduction

	Final Sample (*N* = 618)
Age, M (SD)	27.6(14.5)
Gender, n (%)	
Women	474(76.7)
Men	144(23.3)
Healthcare Background, n (%)	
Practitioner	101(16.3)
General Public	517(83.7)
Neurofeedback Experience of Practitioners[a], n (%)	
Experience	22(3.6)
No Experience	52(8.4)
Age Group Practitioners[b], n (%)	
Younger	49(7.9)
Older	52(8.4)

*Note.*The variables Gender, Healthcare Background, Neurofeedback Experience, and Age Group Practitioners are reported in sample size (n) and percentage (%).

[a]Based on the answer (yes/no) to question Q.13.4 "Do/did you include neurofeedback in your therapies" of the questionnaire. This question was only visible to practitioners who had already heard of Neurofeedback prior to this questionnaire.

[b]Based on median age (*Mdn* = 42yrs) of the practitioner group (Younger <= 42yrs, Older > 42yrs).

Table 2

Behavioral Intention Average Per Variable

	Behavioral Intention Average[a]
Final Sample, M (SD)	67.1(21.0)
Gender, M (SD)	
Women	67.3(21.3)
Men	66.5(20.1)
Healthcare Background, M (SD)	
Practitioner	69.0(21.1)
General Public	66.8(21.0)
Neurofeedback Experience of Practitioners, M (SD)	
Experience	63.6(25.6)
No Experience	71.3(19.7)
Age Group Practitioners, M (SD)	
Younger	67.9(17.7)
Older	70.1(23.9)

Note. The sample size remains the same for each group, as shown in Table 1.

[a]Calculated based on the mean score of questions BI1: Assuming I had access to a neurofeedback companion during my neurofeedback training, I would use it, BI2: I would like to use a neurofeedback companion during my neurofeedback training for the improvement of cognitive abilities, and BI3: If a close relative had this possibility, I would advise them to use a neurofeedback companion during their neurofeedback training.

Main Analyses

Hypothesis 1 – Difference between practitioners and the general public

An independent samples t-test with matched groups was conducted to examine this first hypothesis. Levene's test showed that the assumption of equal variances was not violated, $F(1,122) = 0.01$, $p = .921$, and therefore equal variances were assumed (see Table 4). The Shapiro-Wilk test revealed that the BI scores were not normally distributed for both groups. In the general public group, this assumption was violated: $W = 0.954$, $p = .022$. Normality was not violated by the practitioner group: $W = 0.971$, $p = .158$ (see Table 5). After examining the Q-Q plots (see Figure 5), which showed approximately normal distribution, and based on the equal and relatively large sample size, it was decided to still conduct the independent sample t-test, despite the normality assumption violation, since independent samples t-tests are relatively robust to this violation (Rasch et al., 2007).

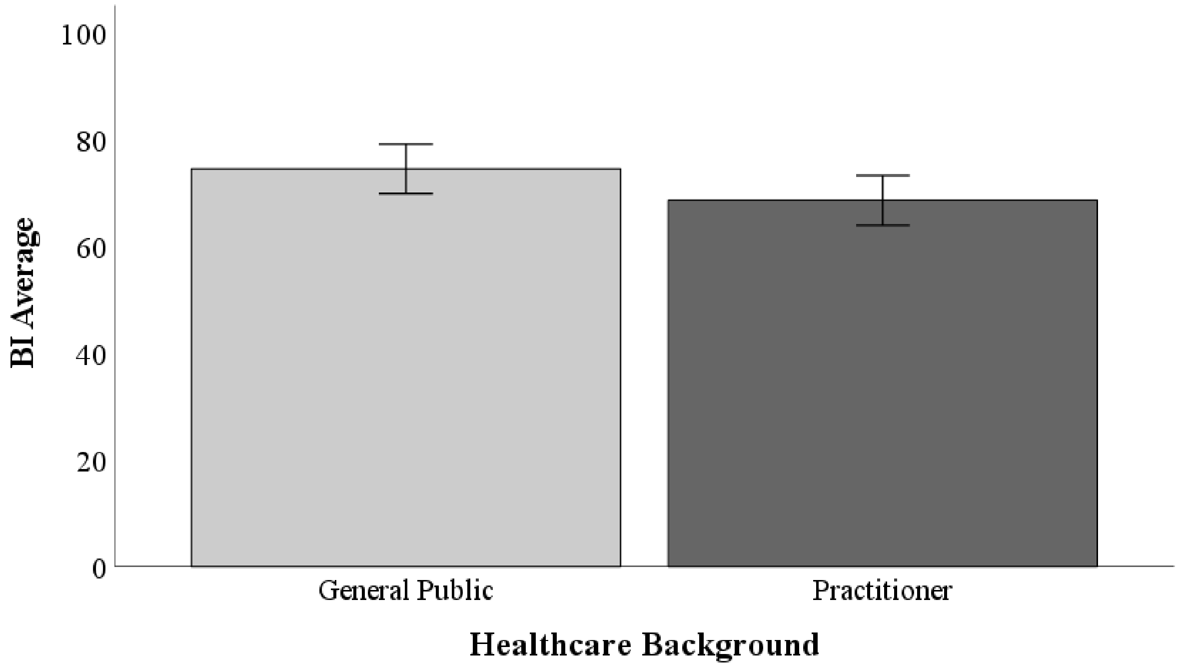
The results showed that the general public reported a higher BI ($M \approx 75$, $SD = 18.2$) than the practitioners ($M \approx 69$, $SD = 18.3$). This difference was statistically significant, $t(122) = 1.8$, $p = .037$ (one-sided). The mean difference between the two groups was approximately -5.91 points on the mean BI score. This indicates that practitioners, on average, scored 5.91 points lower on BI than the general public. Even though the 95% two-sided confidence interval of the difference [-12.41, 0,59] included zero, the one-sided test, which was justified by the directional hypothesis, revealed a statistically significant effect with a small effect size (Cohen's $d = 0.323$). The results[1] are shown below in Figure 1.

---

[1] The full results for the independent samples t-test are reported in Table 6 in Appendix B.

Figure 1

Bar Plot of BI Average by Healthcare Background



*Note.*Error bars represent +/- 2 SEs. BI refers to Behavioral Intention

These findings do not support the hypothesis that practitioners in general will have a higher intention to use a robot-like NF companion than the general public; rather, they suggest the opposite.

Hypothesis 2 – Difference between practitioners with or without NF experience

Initially, it was planned to conduct an independent samples t-test to examine the differences in BI. However, Levene's Test for Equality of Variances revealed that this assumption was violated, $F(1,72) = 4.531$, $p = .037$ (see Table 4). Moreover, the Shapiro-Wilk test indicated a significant deviation from normality for the no NF experience group, $W = .934$, $p = .007$. The NF experience group showed no violation of this assumption, $W = .949$, $p = .307$ (Table 5). Although the Q-Q plots (see Figure 6) were approximately normally distributed, due to the unequal sample sizes, it was decided to conduct a Mann-Whitney U test instead of the independent samples t-test.

The Mann-Whitney U test revealed no statistically significant difference, $U = 477.0$, $Z = -1.12$, $p = .261$, in the behavioral intention to use a robot-like NF companion between practitioners with (*Mdn* = 66.6) and without (*Mdn* = 70) NF experience, with a small effect size ($r = .13$). The findings are reported below in Table 3 and Figure 2.

Table 3

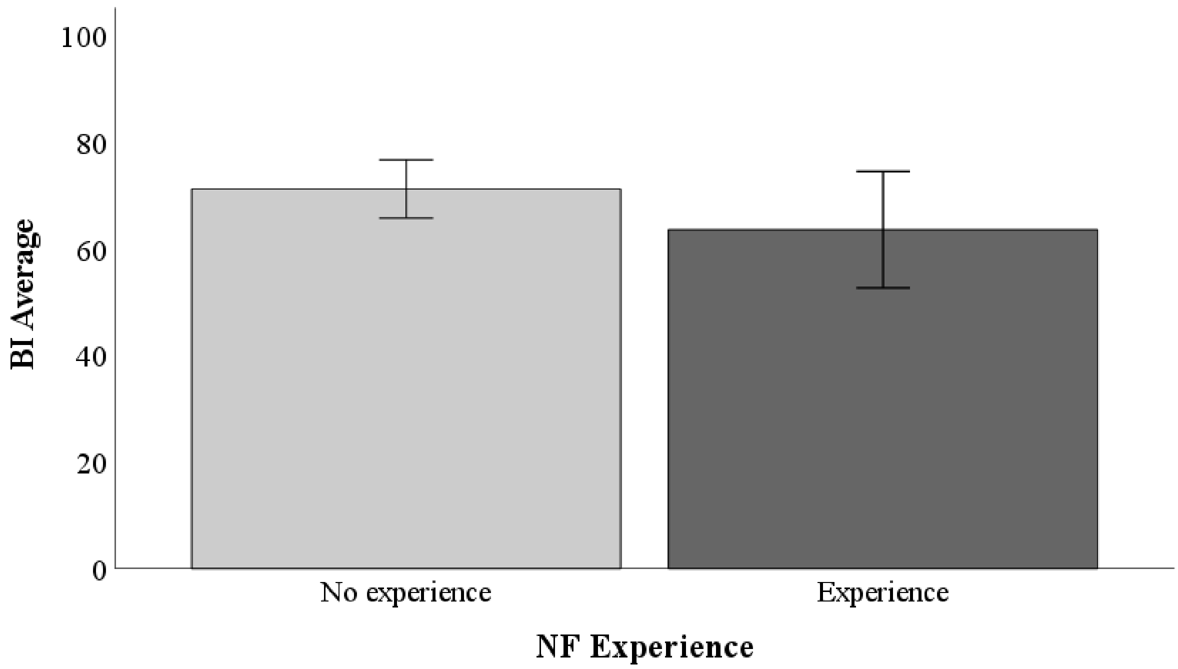
Mann-Whitney U Test Results

	Therapists with Experience		Therapists with no Experience					
	Mean rank	*Sum of ranks*	*Mean rank*	*Sum of ranks*	*U*	*Z*	*p*	*r*
BI Average	33.2	730	39.3	2045.0	477.0	-1.124	.261	.13

Note. BI refers to Behavioral Intention. *U* is the test statistic. *Z* is the z-score of *U*.

Figure 2

Bar Plot of BI Average by NF Experience



Note. Error bars represent +/- 2 SEs. BI refers to Behavioral Intention. NF refers to neurofeedback.

These findings do not support the hypothesis that practitioners with prior NF experience will have a higher intention to use a robot-like NF companion than practitioners without prior experience.

Hypothesis 3 – Difference between ages, NF experience, and gender

A three-way between-subjects ANOVA was conducted to examine this hypothesis. The assumption checks revealed the same violations for the *NF_experience* variable as in the second hypothesis. For the other two variables, Levene's Test for Equality of Variances revealed no violations ($p > .05$) (see Table 4). The Shapiro-Wilk test revealed violations of the normality assumption for both the older, $W = .924$, $p = .006$, and female, $W = .958$, $p =$

.036, groups (see Table 5, Figures 7 and 8). Although these assumptions were violated, it was decided to conduct the 3-way ANOVA[2].

The 3-way ANOVA revealed no statistically significant main effects with very low or low effect sizes for Gender, $F(1,66) = .71$, $p = .402$, partial $\eta^2 = .011$, NF Experience, $F(1,66) = 1.85$, $p = .178$, partial $\eta^2 = .027$, and Age, $F(1,66) = .58$, $p = .45$, partial $\eta^2 = .009$. Additionally, none of the tested interaction effects were statistically significant and had very low effect sizes[3]:

Gender x NF Experience: $F(1,66) = .35$, $p = .556$, partial $\eta^2 = .005$

Gender x Age: $F(1,66) = .31$, $p = .581$, partial $\eta^2 = .005$

Age x NF Experience: $F(1,66) = .013$, $p = .909$, partial $\eta^2 = .000$

Gender x Age x NF Experience: $F(1,66) = .60$, $p = .443$, partial $\eta^2 = .009$

Since none of the main or interaction effects were significant, for exploratory reasons, it was decided to examine the estimated marginal means plots. Separate plots were created based on gender.

The first plot (see Figure 3), based on the male gender group, showed that male practitioners with previous NF experience demonstrated a big increase in BI from younger to older. In contrast, male practitioners without prior NF experience only showed a minimal increase in BI from younger to older.

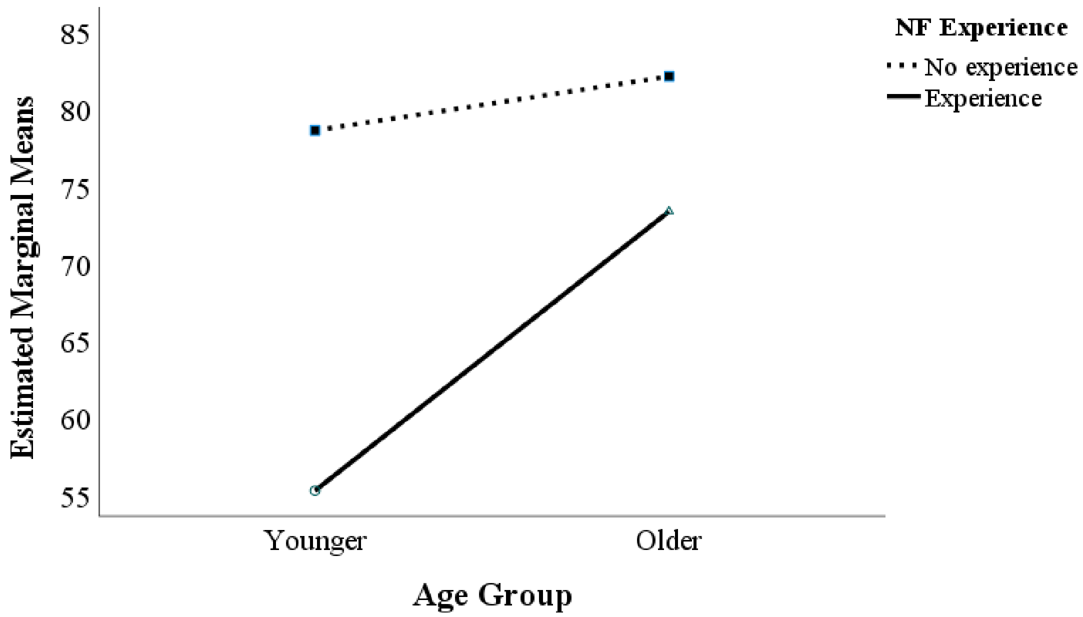
---

[2] Although a Generalized Linear Model would be better suited for this analysis, this statistical technique is not part of the curriculum. Therefore, results must be interpreted with caution.

[3] The full results are reported in Table 7 of Appendix B.

Figure 3

Estimated Marginal Means Plot of BI Average at Gender = Male

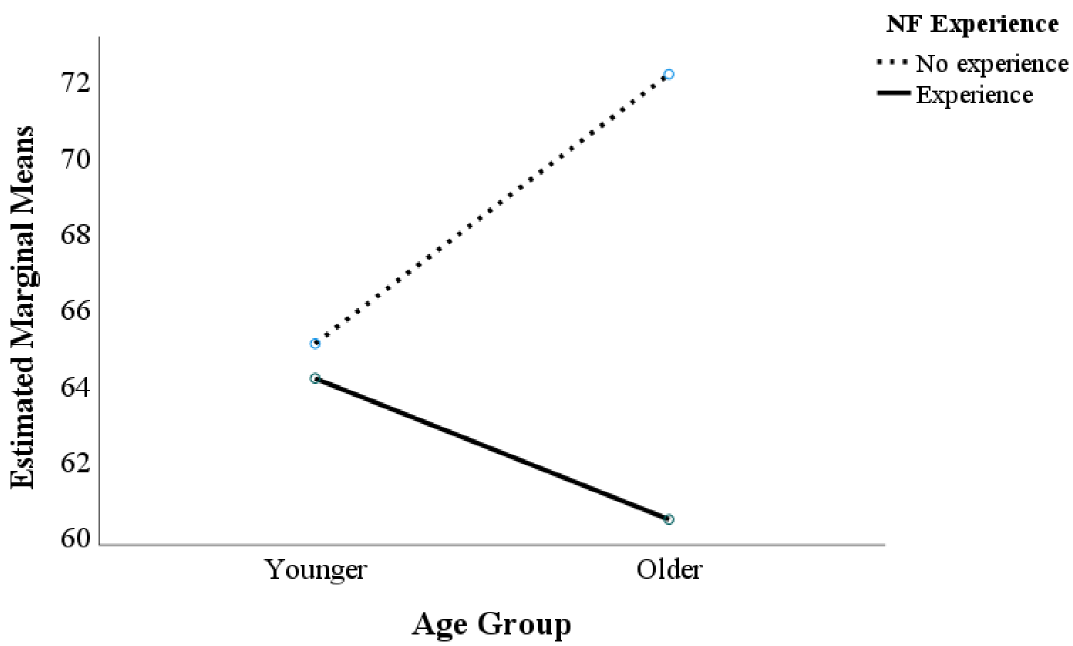


*Note.*The age groups are based on the median age (*Mdn* = 42yrs) of the practitioner group (Younger <= 42yrs, Older > 42yrs).

For the second plot (see Figure 4), based on the female gender group, the pattern is reversed. Female practitioners with previous NF experience demonstrated a drop in BI from younger to older. Moreover, female practitioners without prior NF experience showed an increase in intention from younger to older.

Figure 4

Estimated Marginal Means Plot of BI Average at Gender = Female



Note. The age groups are based on the median age (*Mdn* = 42yrs) of the practitioner group (Younger <= 42yrs, Older > 42yrs).

Overall, the results do not show support for the hypothesis that younger practitioners will have a higher intention to use a robot-like NF companion than older practitioners and that this relationship will be stronger among those with neurofeedback experience compared to those without, which will be even more pronounced by gender.

Discussion

Implementing new technologies into current clinical practice is a lengthy and time-consuming process (Balas & Boren, 2000). To assess whether a new technology is worth implementing, it is crucial to assess its acceptability by practitioners (Liu et al., 2014). Thus, this study aimed to investigate whether there is a difference in the BI to use a robot-like NF companion for NF training between practitioners already working with NF and those who are not, compared to the general public. In the following section, the results of the analyses above will be discussed and contextualized, while also trying to answer the question of how to implement the robot-like NF companion into clinical practice.

Hypothesis 1 – Differences Between Practitioners and the General Public

The results related to the first hypothesis indicated that practitioners had a significantly lower BI to use a robot-like NF companion than the general public. This contradicted the hypothesis. A reason for this could be that the general public may lack domain-specific knowledge regarding NF and therefore cannot accurately assess the robot-like NF companion and show less apprehension to use it, which results in a higher BI score than practitioners. Domain-specific knowledge refers to the fact that practitioners might know more about NF itself, the mechanisms at play, and how the type of feedback provided to the participant can help their performance and progression compared to the general public. According to Chi (2006), the general public would be classed as "Naïve", with no knowledge of the domain, and practitioners as "Experts" with extensive knowledge derived from experience. Additionally, it is important to contextualize the aforementioned result. It should be noted that the scores from both groups do not differ dramatically on average, and both score relatively high on BI, ranging from 68.9 to 74.6 points out of a possible 100. This suggests a generally positive attitude of both groups towards using the robot-like NF companion. Interestingly, when looking at the unmatched full sample, practitioners show a higher average BI score than the general public, 69 and 66.8, respectively (see Table 2), which is more in line with the original hypothesis. This difference may partially be explained by the drastically different sample sizes between the matched and unmatched groups – 62 cases per group in the matched condition versus 101 practitioners and 517 general public cases in the unmatched condition.

Larger and more heterogeneous samples, like the unmatched condition, may show a greater variability in individual attitudes than the matched groups, which could lower the BI

(Frankot et al., 2023). This is especially visible for the general public group. Additionally, statistical matching procedures are designed to balance the distribution of possible covariates in the matched groups, in this case, age and gender, to reduce variability (Stuart, 2010). However, the matching procedure used in this paper removed cases that did not have an exact age and gender match. This could have led to the matched groups no longer representing the broader population from which the sample was drawn, reducing external validity and representativeness of the matched sample (Stuart, 2010). Moreover, the fact that both groups scored high on BI suggests that there may not be an actual, meaningful difference between the groups in the real world, even though it was statistically significant (McShane et al., 2019). Moreover, Cohen's *d* indicated only a small effect, further supporting this argumentation. Lastly, the reversal of BI scores between the two conditions resembles Simpson's Paradox (Wagner, 1982). It illustrates that group trends can be reversed when data are combined across different groups, without considering subgroup differences. Although this is not a textbook case of Simpson's Paradox, the results show that it is good to be aware of how sample composition can have an impact on group-level outcomes.

Hypothesis 2 – Differences Between Practitioners with and without Prior NF Experience

The analysis pertaining to the second hypothesis revealed that both groups did not differ significantly in their BI, which contradicts the hypothesis. This may indicate that prior NF experience alone may not have a significant effect on the intention to use or implement the NF companion. However, the unequal sample sizes (22 with experience and 52 without) might have reduced the ability to detect a statistically significant effect. It is also possible that the other aspects of acceptability (perceived usefulness and perceived ease of use) are better predictors for the adoption of a robot-like NF companion by practitioners, which is in line with prior research on the adoption of new technologies by therapists (Grevet et al., 2023; Liu et al., 2014). Additionally, as Grevet et al. (2023) pointed out, previous experience acts as a moderator for the relationship between BI and technology adoption; the more experience a user has with the new technology via previous experience, the less influence BI has on the actual use of the new technology. Therefore, it is possible that the previous NF experience that some of the practitioners had did not influence their BI to use the robot-like NF companion. This may explain why we could not show a significant difference between the two groups.

Hypothesis 3 – Differences Between Ages, Genders, and Experience Levels

The results related to the third hypothesis revealed no statistically significant main or interaction effects of age, gender, or NF experience. This suggests that these demographic factors and prior NF experience may not significantly affect practitioners' BI to use the new robot-like NF companion in this analysis. To ultimately determine their effects, the calculation of Bayesian statistics would be needed. Bayes Factor assesses the data in favor of the null hypothesis, which is not possible in traditional frequentist statistics, and can provide a better interpretation of non-significant results (Dienes, 2014) It is also possible that the small and unequal sample sizes per group also reduced the ability to detect a significant effect, as well as raised the possibility of Type II errors. These results differ from previous research on new technology adoption that underlines age and gender differences as important factors (Anderson & Perrin, 2017; Berkowsky et al., 2018; Venkatesh et al., 2000; Yau & Cheng, 2012). Additionally, the findings conflict with both versions of the Unified Theory of Acceptance and Use of Technology, which identified age, gender, and prior experience as important moderators for BI (Venkatesh et al., 2003; Venkatesh et al., 2012). Therefore, age and gender may not play an important role in this particular intervention. Moreover, it is possible that other factors, such as organizational environment, perceived usefulness, and personal characteristics, are more important to practitioners when implementing a new technology and therefore override BI, which would be in line with the Technology

Acceptance Model and the Unified Theory of Acceptance and Use of Technology 2 (Venkatesh et al., 2012; Yarbrough & Smith, 2007). An organizational environment can also impact acceptability as it can offer facilitating conditions during the implementation of new technologies, as training and support are usually freely available within organizations (Venkatesh et al., 2012). Thus, it is possible that our robot-like NF companion does not fit into the definition of a new technology of those models, as it may be too specific for them.

Examining the interaction plots, one could see that in the male gender group, BI increased a lot from the younger to older practitioners with prior NF experience. In contrast, for those with no experience, the increase was only slight. The pattern was reversed for the female group. Female practitioners without experience demonstrated an increase in BI from younger to older, while the BI dropped slightly from younger to older for female practitioners with prior experience. These interaction plots could suggest possible gender-based differences in how age and prior NF experience interact to influence BI to implement the robot-like NF companion. These findings align with previous research (Anderson & Perrin, 2017; Berkowsky et al., 2018; Venkatesh et al., 2000; Yau & Cheng, 2012). For example, Venkatesh et al. (2000) found that gender plays an important role in the initial decision-making process of new technology adoption and subsequent usage behavior, while Anderson and Perrin (2017) found that technology adoption is climbing amongst older, well-educated adults. Additionally, Yau & Cheng (2012) discovered that male students were more confident in using technology, which you can also see in the difference in BI between younger male and female practitioners without experience. These findings, once again, highlight the importance of not only relying on significance testing when interpreting data (McShane et al., 2019).

Bringing the Robot-Like NF Companion to Clinical Usage

The central concern that was outlined at the beginning of this paper was the so-called "research-to-practice gap" that is present in clinical neuropsychology. In the context of implementing a new robot-like NF companion, this gap represents a substantial hurdle. However, the results of the analyses above can be used to bring the robot-like NF companion into practice. Across all analyses, the BI average was above the score of 60 out of 100, which indicates a generally favorable attitude towards the robot-like NF companion and a willingness to use it. To fully implement it, the next steps should include further testing of the robot-like NF companion in practice with target populations that can benefit from it, as well as practitioners, to allow them to get familiar with the technology and enhance their knowledge and experience. This is important since Proctor et al. (2011), in their quest to clarify the language used in implementation research, defined acceptability as a changing concept regarding experience and knowledge. Mitchell et al. (2022) have proven this with their study on the acceptability of a training program for practitioners working with precision medicine. They used small, guided focus groups with relevant practitioners to allow them to enhance their knowledge of the intervention, outline barriers to implementing the intervention, and discuss possible future training programs aimed at implementation (strengths, risks, challenges) (Mitchell et al., 2022). Focus groups are a commonly used method to discuss a specific topic in depth with relevant people and draw from these people's personal experiences, beliefs, and concerns (Nyumba et al., 2018). After the focus groups, Mitchell et al. (2022) found that the initially low acceptance of the training program had increased. Implementing such focus groups for relevant practitioners (practitioners with prior NF experience) regarding the robot-like NF companion could help to increase its acceptance among the different age groups and genders, as outlined by the possible interaction trends of the third hypothesis. It is imperative to translate our research findings into clear and concise summaries that can actually be used by practitioners to aid the implementation of the robot-like NF companion. These summaries could then be used in the focus groups to enhance knowledge, understanding, and therefore acceptability of the robot-like NF companion. It is important to translate our findings into usable formats, as implementation knowledge is not taught to practitioners (Ovretveit et al., 2017), and knowledge of implementation science still largely belongs to researchers and not practitioners (Westerlund et al., 2019).

Strengths and Limitations

To the best of my knowledge, this is the first study to examine the relationship between BI scores to use a robot-like NF companion by practitioners and the general public. Due to this, the interpretations should be taken with a grain of salt. Additionally, several strengths should be considered**.** Firstly, the overall sample size is large and international. This allowed us to conduct subgroup analyses and gain a robust and diverse insight into the opinions of groups who may come into contact with the robot-like NF companion in the future, regarding its implementation. It also allowed us to conduct interaction analyses, which would not have been possible with smaller sample sizes. Additionally, because the questionnaire was available in multiple languages and online, it allowed us to reach a diverse group of people from many different countries across Europe and even outside of it to gain a nuanced understanding of their opinions. Moreover, the questionnaire covers the main aspects that are relevant for the acceptability of a new technology and further demographic and personality factors that may influence the acceptability. This allows us to draw multifaceted conclusions based on their answers with regard to the future steps of the implementation process.

Additionally, several limitations should be considered. First, the study lacked practitioners, especially ones with prior NF experience, who filled out the questionnaire. Moreover, women and younger people were overrepresented in our sample, while men and older people were underrepresented. This led to very unequal and small sample sizes for each of the subgroups and may impact the generalizability of the results to the broader population. To overcome this, future studies should use larger, more representative samples concerning the aforementioned groups. Researchers should more directly try to recruit these target demographics. This would also help to investigate subgroup effects and interaction patterns more reliably.

Second, we used a convenience sampling method. This can further affect the generalizability of our findings by introducing selection bias. To address this, stratified sampling could be employed. This means dividing the population into relevant subgroups (e.g., healthcare background, genders, etc.) and randomly sampling within these groups.

Lastly, the questionnaire that was used was relatively long, time-intensive, and complex, with the acceptability portion of it being the second-to-last section. Due to this, it is possible that participants were affected by respondent fatigue, especially in the later parts of the questionnaire, which may have affected the quality of their responses. It is possible that participants were not carefully reading the questions or considering their responses at this point. Because of the complexity, some participants may have struggled to understand certain key concepts, like what NF is and what the robot-like NF companion is trying to improve, even with the comprehension check section in place. If people misunderstood the role and nature of the robot-like NF companion, it is possible that they falsely interpreted the acceptability portion of the study. This could also have an impact on the validity of the data. To mitigate this, the overall questionnaire design could be improved by shortening, simplifying, and rearranging the questions. One possible starting point could be the shortening of the personality section, as it is one of the longest, with 44 questions. The shortening and simplification could help with respondent fatigue, while putting the acceptability-related questions earlier could ensure proper understanding of these concepts.

Future Research Recommendations

Based on the results, some future research recommendations can be discerned. Firstly, employing qualitative data, like interviews and focus groups, could help to understand certain attitudes and concerns of practitioners regarding the robot-like NF companion better than solely relying on the questionnaire, since people can give more nuanced answers this way. This is especially important since the questionnaire did not give participants the option to express concerns (outside of scales) at all. Possible concerns could include ethics and privacy of the patient data, possible interference with therapeutic alliance, or additional training and workload for practitioners. Some possible questions to ask are listed below:

1. *What concerns, if any, do you have about the implementation of the robot-like NF companion into your clinical practice?*
2. *How well would the robot-like NF fit into your current clinical practice?*
3. *What would make you more willing to implement such a tool into clinical practice?*
4. *Do you think you would benefit from future training programs or supporting material designed to help you integrate the robot-like NF companion into practice?*

Secondly, future studies should take into account organizational context when assessing the acceptability of practitioners, since this aspect, while being part of different acceptability models, is not assessed in the questionnaire. To do so, the Implementation Climate Scale, which was developed by Ehrhart et al. (2014), could be added. This brief scale (18 items) assesses how strategic the organizational climate is for implementing evidence-based practices based on employees' perceptions. It covers six dimensions, which include educational support and recognition for the evidence-based practice that may be implemented. The scale can help to understand the organizational context's role in possibly implementing the robot-like NF companion. This part of the questionnaire should only be visible to relevant subgroups, i.e., practitioners who work with NF.

Lastly, all the analyses should be redone in future studies with larger and more equal sample sizes to see whether the lack of findings can be attributed to this shortcoming.

Conclusion

Overall, our robot-like NF companion was generally viewed favorably, and it should now be tested during actual NF trials. Additionally, practitioner involvement in future feasibility studies should be considered. With their involvement and input, a straightforward integration into clinical routines (e.g., how to set it up and when it can be helpful) and development of training sessions and support resources (e.g., guidelines on how to work with it and introduce it to participants) may be established. This step could also help to bring the robot-like NF companion into clinical practice.

References

Ahuna, J. K., Becker, K. D., & Chorpita, B. F. (2023). Predicting Therapists' Intentions to Use Innovations: Comparing the Role of Individual, Organizational, and Innovation Characteristics. *Administration and Policy in Mental Health and Mental Health Services Research*, *50*(6), 946–965. https://doi.org/10.1007/s10488-023-01295-6

Anderson, M., & Perrin, A. (2017, May 17). Tech Adoption Climbs Among Older Adults. *Pew Research Center*. https://assets.pewresearch.org/wp-content/uploads/sites/14/2017/05/16170850/PI_2017.05.17_Older-Americans-Tech_FINAL.pdf

Assaker, G. (2019). Age and gender differences in online travel reviews and user-generated-content (UGC) adoption: extending the technology acceptance model (TAM) with credibility theory. *Journal of Hospitality Marketing & Management*, *29*(4), 1–22. https://doi.org/10.1080/19368623.2019.1653807

Balas, E. A., & Boren, S. A. (2000). Managing Clinical Knowledge for Health Care Improvement. *Yearbook of Medical Informatics*, 65–70.

Beidas, R., Skriner, L., Adams, D., Wolk, C. B., Stewart, R. E., Becker-Haimes, E., Williams, N., Maddox, B., Rubin, R., Weaver, S., Evans, A., Mandell, D., & Marcus, S. C. (2017). The relationship between consumer, clinician, and organizational characteristics and use of evidence-based and non-evidence-based therapy strategies in a public mental health system. *Behaviour Research and Therapy*, *99*, 1–10. https://doi.org/10.1016/j.brat.2017.08.011

Berkowsky, R., Sharit, J., & Czaja, S. (2018). Factors Predicting Decisions About Technology Adoption Among Older Adults. *Innovation in Aging*, *1*(3), 1–12. https://doi.org/10.1093/geroni/igy002

Bowen, D. J., Kreuter, M., Spring, B., Cofta-Woerpel, L., Linnan, L., Weiner, D., Bakken, S., Kaplan, C. P., Squiers, L., Fabrizio, C., & Fernandez, M. (2009). How we design feasibility studies. *American Journal of Preventive Medicine*, *36*(5), 452–457. https://doi.org/10.1016/j.amepre.2009.02.002

Chi, M. T. H. (2006). Two Approaches to the Study of Experts' Characteristics. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge Handbook of Expertise and Expert Performance* (pp. 21–30). Cambridge University Press. https://doi.org/10.1017/CBO9780511816796.002

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. https://doi.org/10.1016/s0360-1315(02)00130-6

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00781

Ehrhart, M. G., Aarons, G. A., & Farahnak, L. R. (2014). Assessing the organizational context for EBP implementation: the development and validity testing of the Implementation Climate Scale (ICS). *Implementation Science*, *9*(1). https://doi.org/10.1186/s13012-014-0157-1

Enriquez-Geppert, S., Huster, R. J., Figge, C., & Herrmann, C. S. (2014). Self-regulation of frontal-midline theta facilitates memory updating and mental set shifting. *Frontiers in Behavioral Neuroscience*, *8*. https://doi.org/10.3389/fnbeh.2014.00420

Frankot, M. A., Young, M. E., & Cole Vonder Haar. (2023). Understanding Individual Subject Differences through Large Behavioral Datasets: Analytical and Statistical Considerations. *Perspectives on Behavior Science*, *47*(1), 225–250. https://doi.org/10.1007/s40614-023-00388-9

Grevet, E., Forge, K., Tadiello, S., Izac, M., Amadieu, F., Brunel, L., Pillette, L., Py, J., Gasq, D., & 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*. https://doi.org/10.3389/fnrgo.2022.1082901

Hammond, D. C. (2007). What Is Neurofeedback? *Journal of Neurotherapy*, *10*(4), 25–36. https://doi.org/10.1300/j184v10n04_04

Holden, R. J., & Karsh, B.-T. (2010). The Technology Acceptance Model: Its past and Its Future in Health Care. *Journal of Biomedical Informatics*, *43*(1), 159–172. https://doi.org/10.1016/j.jbi.2009.07.002

Izuma, K., Saito, D. N., & Sadato, N. (2008). Processing of Social and Monetary Rewards in the Human Striatum. *Neuron*, *58*(2), 284–294. https://doi.org/10.1016/j.neuron.2008.03.020

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). Big five inventory. *Journal of Personality and Social Psychology*, *1*(1). https://doi.org/10.1037/t07550-000

Keegan, R. J., Cotterill, S., Woolway, T., Appaneal, R., & Hutter, V. (2017). Strategies for bridging the research-practice "gap" in sport and exercise psychology. *Revista de Psicologia Del Deporte*, *26*(4), 75–80.

Liu, L., Cruz, A., Rincon, A., Buttar, V., Ranson, Q., & Goertzen, D. (2014). What factors determine therapists' acceptance of new technologies for rehabilitation -a study using the Unified Theory of Acceptance and Use of Technology (UTAUT). *Disability and Rehabilitation*, 1–9. https://doi.org/10.3109/09638288.2014.923529

Marzbani, H., Marateb, H., & Mansourian, M. (2016). Methodological Note: Neurofeedback: A Comprehen- sive Review on System Design, Methodology and Clini- cal Applications. *Basic and Clinical Neuroscience*, *7*(2). https://doi.org/10.15412/J.BCN.03070208

Mathiak, K. A., Alawi, E. M., Koush, Y., Dyck, M., Cordes, J. S., Gaber, T. J., Zepf, F. D., Palomero-Gallagher, N., Sarkheil, P., Bergert, S., Zvyagintsev, M., & Mathiak, K. (2015). Social reward improves the voluntary control over localized brain activity in fMRI-based neurofeedback training. *Frontiers in Behavioral Neuroscience*, *9*. https://doi.org/10.3389/fnbeh.2015.00136

Mathiak, K. A., Koush, Y., Dyck, M., Gaber, T. J., Alawi, E., Zepf, F. D., Zvyagintsev, M., & Mathiak, K. (2010). Social reinforcement can regulate localized brain activity. *European Archives of Psychiatry and Clinical Neuroscience*, *260*(S2), 132–136. https://doi.org/10.1007/s00406-010-0135-9

Matsuzaki, Y., Nouchi, R., Sakaki, K., Dinet, J., & Kawashima, R. (2023). The Effect of Cognitive Training with Neurofeedback on Cognitive Function in Healthy Adults: A Systematic Review and Meta-Analysis. *Healthcare*, *11*(6), 843. https://doi.org/10.3390/healthcare11060843

McShane, B. B., Gal, D., Gelman, A., Robert, C., & Tackett, J. L. (2019). Abandon Statistical Significance. *The American Statistician*, *73*(sup1), 235–245. https://doi.org/10.1080/00031305.2018.1527253

Michaelis, J. E., & Mutlu, B. (2017). Someone to Read with. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3025453.3025499

Mitchell, S., Jaccard, E., Schmitz, F. M., von Känel, E., Collombet, P., Cornuz, J., Waeber, G., Guessous, I., & Guttormsen, S. (2022). Investigating acceptability of a training programme in precision medicine for frontline healthcare professionals: a mixed methods study. *BMC Medical Education*, *22*(1). https://doi.org/10.1186/s12909-022-03613-2

Naismith, S. L., Michaelian, J. C., Low, L.-F., Arsenova, V., Mehrani, I., Fyfe, K., Kochan, N. A., Kurrle, S. E., Rowe, C., & Sachdev, P. S. (2022). Characterising Australian memory clinics: current practice and service needs informing national service guidelines. *BMC Geriatrics*, *22*(1). https://doi.org/10.1186/s12877-022-03253-7

Norman, D. A. (2010). The research-practice gap: The Need for Translational Developers. *Interactions*, *17*(4), 9. https://doi.org/10.1145/1806491.1806494

Nyumba, T., Wilson, K., Derrick, C. J., & Mukherjee, N. (2018). The Use of Focus Group Discussion methodology: Insights from Two Decades of Application in Conservation. *Methods in Ecology and Evolution*, *9*(1), 20–32. Wiley. https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.12860

Olswang, L. B., & Prelock, P. A. (2015). Bridging the Gap Between Research and Practice: Implementation Science. *Journal of Speech, Language, and Hearing Research*, *58*(6). https://doi.org/10.1044/2015_jslhr-l-14-0305

Ovretveit, J., Mittman, B., Rubenstein, L., & Ganz, D. A. (2017). Using implementation tools to design and conduct quality improvement projects for faster and more effective improvement. *International Journal of Health Care Quality Assurance*, *30*(8), 755–768. https://doi.org/10.1108/ijhcqa-01-2017-0019

Pike, K. E., Li, L., Naismith, S. L., Bahar-Fuchs, A., Lee, A., Mehrani, I., Bentvelzen, A., Lautenschlager, N. T., O'Connell, M. E., Blackberry, I., & Mowszowski, L. (2024). Implementation of Cognitive (Neuropsychological) Interventions for Older Adults in Clinical or Community Settings: A Scoping Review. *Neuropsychology Review*. https://doi.org/10.1007/s11065-024-09650-6

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. https://doi.org/10.1016/j.ijhcs.2019.102380

Pindi, P., Houenou, J., Piguet, C., & Favre, P. (2022). Real-time fMRI neurofeedback as a new treatment for psychiatric disorders: A meta-analysis. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *119*, 110605. https://doi.org/10.1016/j.pnpbp.2022.110605

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., Griffey, R., & Hensley, M. (2011). Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health and Mental Health Services Research*, *38*(2), 65–76. https://doi.org/10.1007/s10488-010-0319-7

Rasch, D., Teuscher, F., & Guiard, V. (2007). How robust are tests for two independent samples? *Journal of Statistical Planning and Inference*, *137*(8), 2706–2720. https://doi.org/10.1016/j.jspi.2006.04.011

Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research*, *78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Smit, D., Dapor, C., Janneke Koerts, Tucha, O. M., Huster, R. J., & Enriquez-Geppert, S. (2023). Long-term improvements in executive functions after frontal-midline theta neurofeedback in a (sub)clinical group. *Frontiers in Human Neuroscience*, *17*. https://doi.org/10.3389/fnhum.2023.1163380

Stuart, E. A. (2010). Matching Methods for Causal Inference: A Review and a Look Forward. *Statistical Science*, *25*(1), 1–21. https://doi.org/10.1214/09-sts313

Van Doren, J., Arns, M., Heinrich, H., Vollebregt, M. A., Strehl, U., & K. Loo, S. (2018). Sustained effects of neurofeedback in ADHD: a systematic review and meta-analysis. *European Child & Adolescent Psychiatry*, *28*(3), 293–305. https://doi.org/10.1007/s00787-018-1121-4

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information technology: toward a Unified View. *MIS Quarterly*, *27*(3), 425–478. https://doi.org/10.2307/30036540

Venkatesh, V., Morris, M., & Ackerman, P. (2000). A Longitudinal Field Investigation of Gender Differences in Individual Technology Adoption Decision-Making Processes. *Organizational Behavior and Human Decision Processes*, *83*(1), 33–60. https://doi.org/10.1006/obhd.2000.2896

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer Acceptance and Use of Information technology: Extending the Unified Theory of Acceptance and Use of Technology. *MIS Quarterly*, *36*(1), 157–178. https://doi.org/10.2307/41410412

Viviani, G., & Vallesi, A. (2021). EEG-neurofeedback and executive function enhancement in healthy adults: A systematic review. *Psychophysiology*, *58*(9). https://doi.org/10.1111/psyp.13874

Wagner, C. H. (1982). Simpson's Paradox in Real Life. *The American Statistician*, *36*(1), 46–48. https://doi.org/10.1080/00031305.1982.10482778

Westerlund, A., Nilsen, P., & Sundberg, L. (2019). Implementation of Implementation Science Knowledge: The Research-Practice Gap Paradox. *Worldviews on Evidence-Based Nursing*, *16*(5), 332–334. https://doi.org/10.1111/wvn.12403

Williams, N. J., & Beidas, R. S. (2018). Annual Research Review: The state of implementation science in child psychology and psychiatry: a review and suggestions to advance the field. *Journal of Child Psychology and Psychiatry*, *60*(4), 430–450. https://doi.org/10.1111/jcpp.12960

Yarbrough, A. K., & Smith, T. B. (2007). Technology Acceptance among Physicians. *Medical Care Research and Review*, *64*(6), 650–672. https://doi.org/10.1177/1077558707305942

Yau, H. K., & Cheng, A. L. F. (2012). Gender Difference of Confidence in Using Technology for Learning. *Journal of Technology Studies*, *38*(2), 74–79. https://doi.org/10.21061/jots.v38i2.a.2

Appendix A – Assumptions

Table 4

Levene's Test of Homogeneity of Variance for all analyses

	Levene Statistic	*df1*	*df2*	*p*
Healthcare Background	.010	1	122	.921
Neurofeedback Experience	4.531	1	72	.037
Age Group	.744	1	72	.391
Gender	.548	1	72	.461

Note. Based on mean

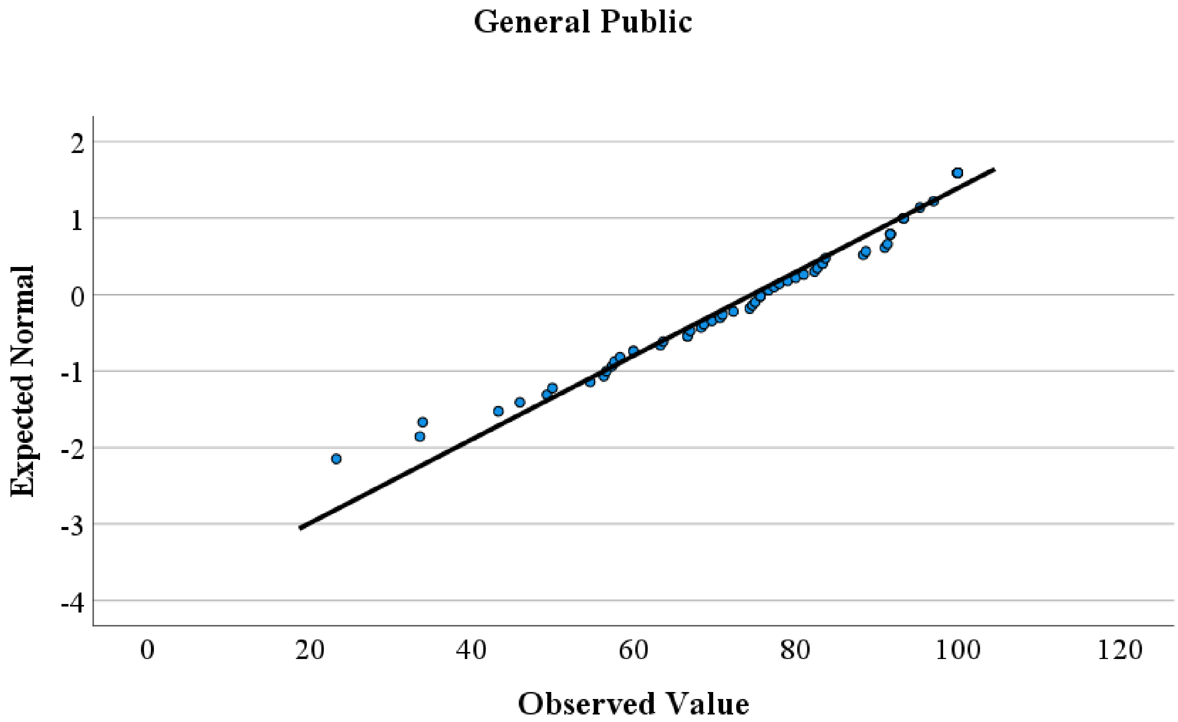
Table 5

Shapiro-Wilk Test of Normality for all analyses

		Statistic	*df*	*p*
Healthcare Background	Practitioner	.971	62	.158
	General Public	.954	62	.022
Neurofeedback Experience	Experience	.949	22	.307
	No Experience	.934	52	.007
Age Group	Younger	.942	30	.103
	Older	.924	44	.006
Gender	Female	.958	60	.036
	Male	.954	14	.618

Figure 5

Q-Q Plots for Behavioral Intention Average by Healthcare Background



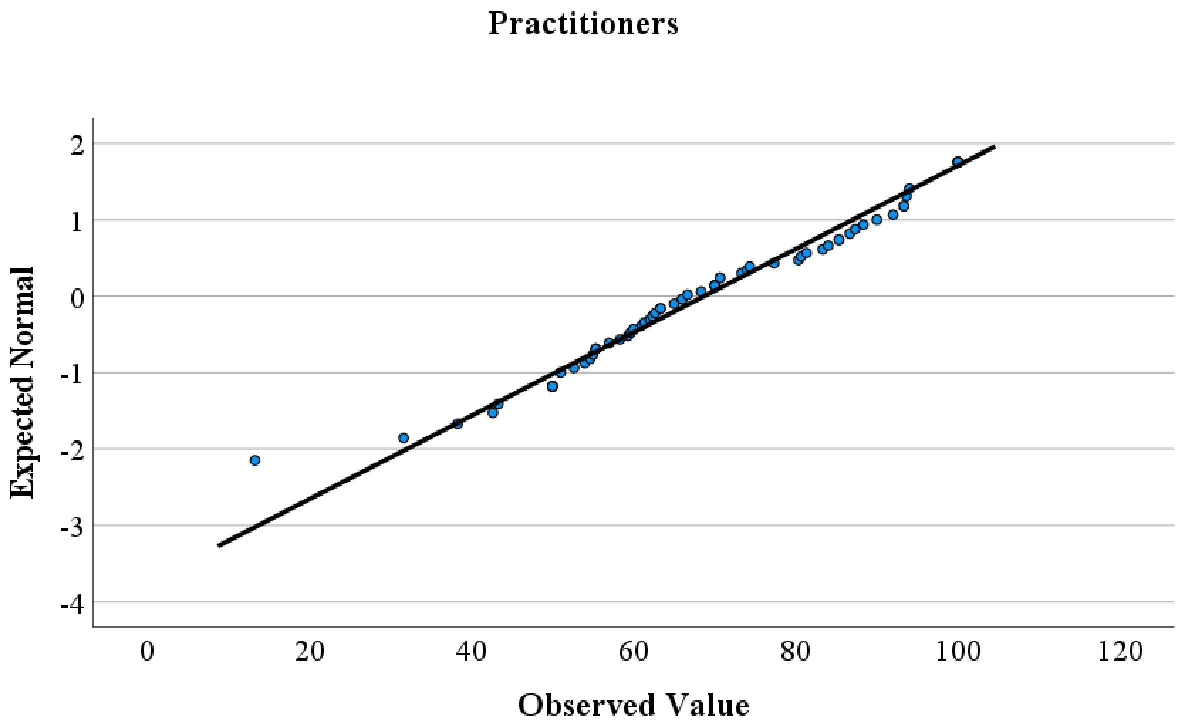


Figure 6

Q-Q Plots for Behavioral Intention Average by Neurofeedback Experience

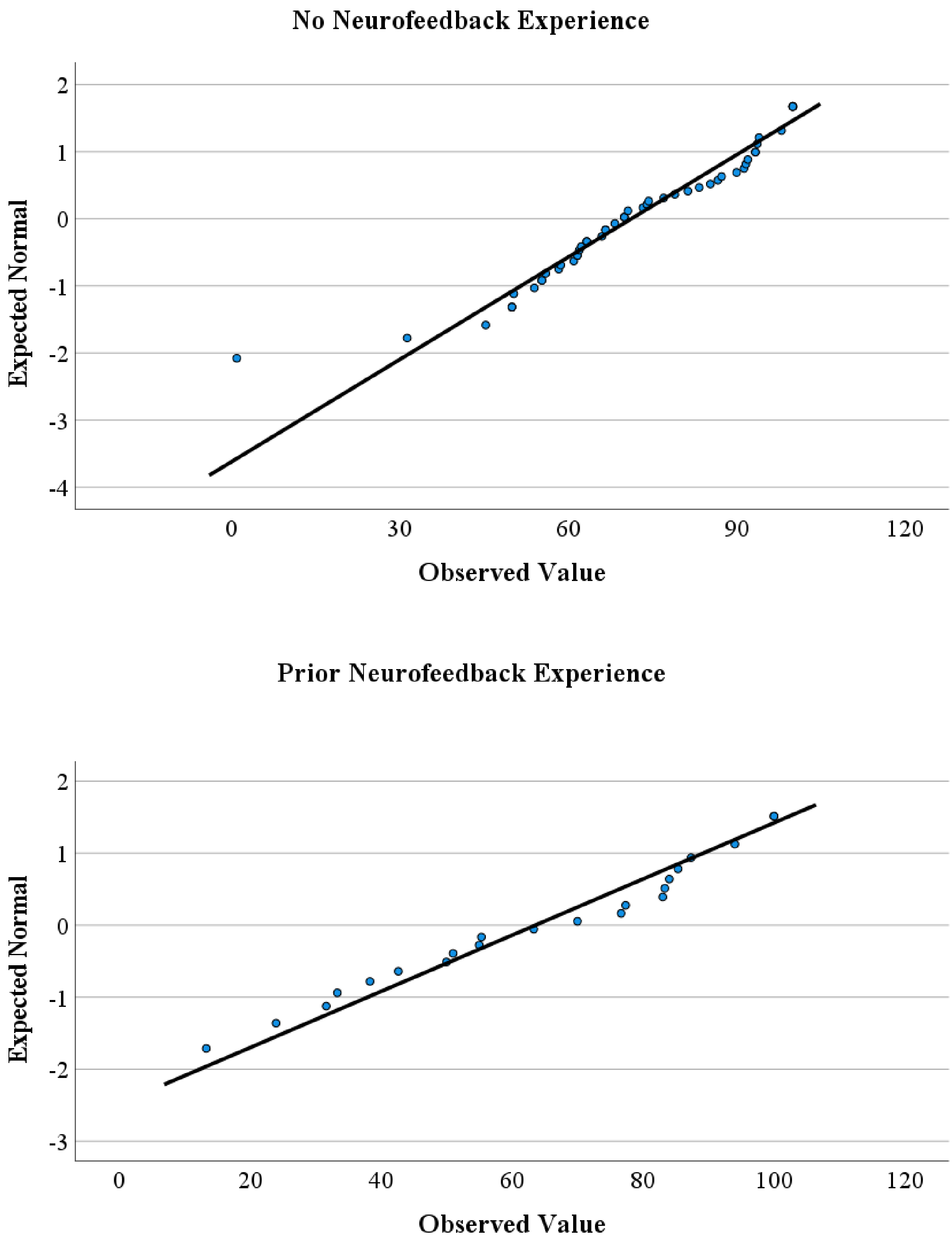
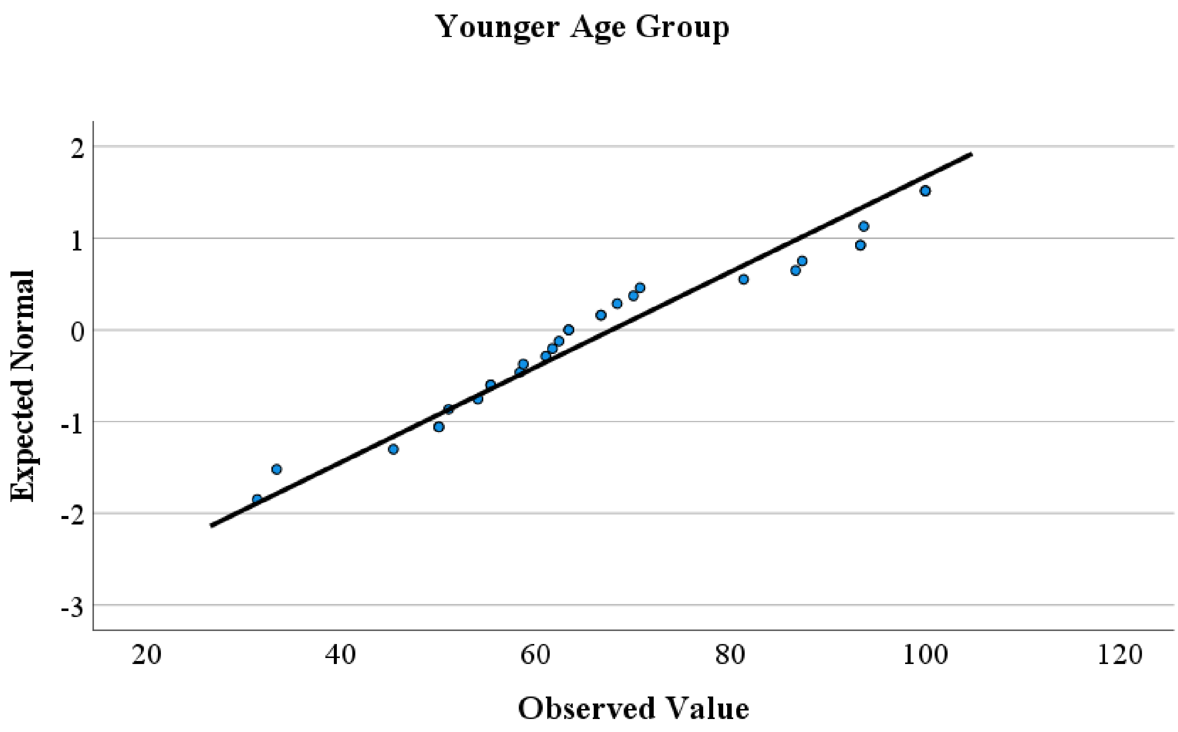


Figure 7

Q-Q Plots for Behavioral Intention Average by Age Groups



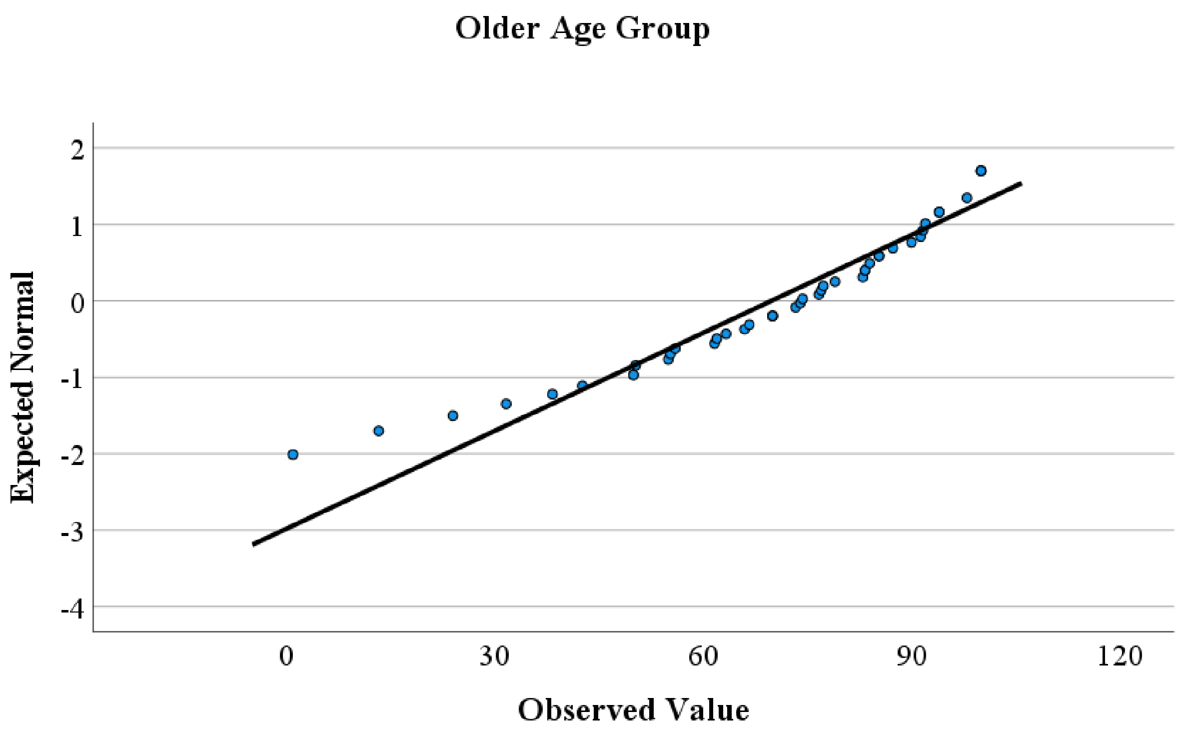
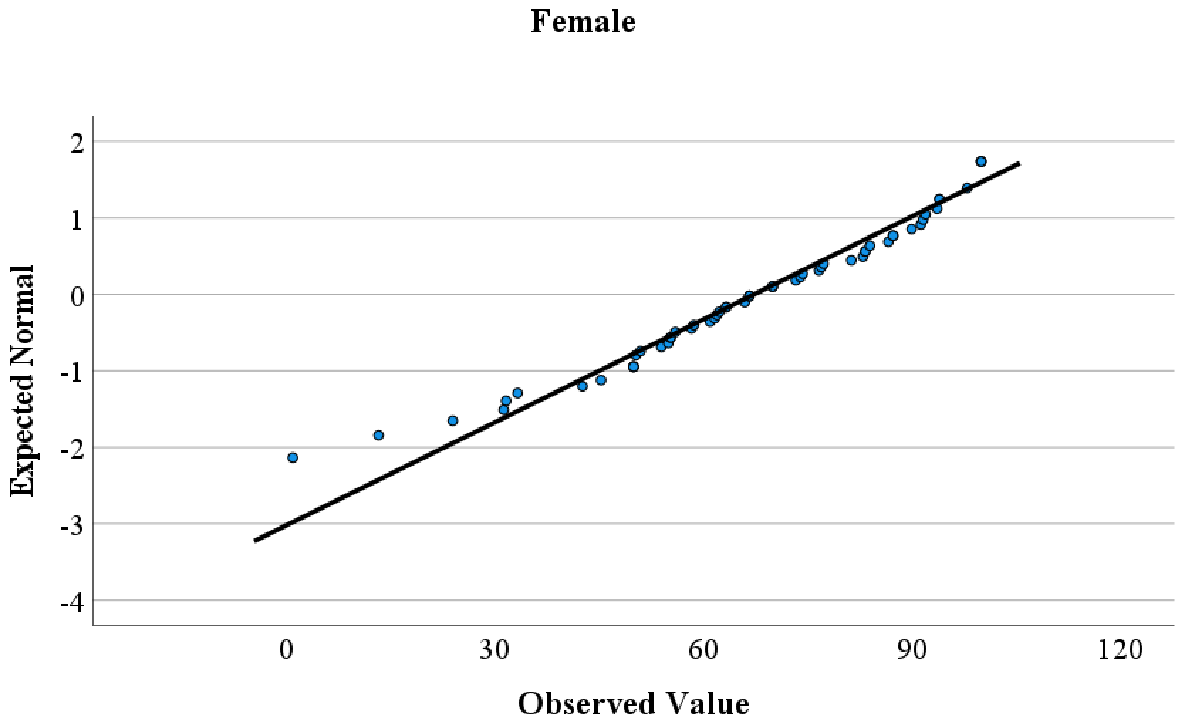
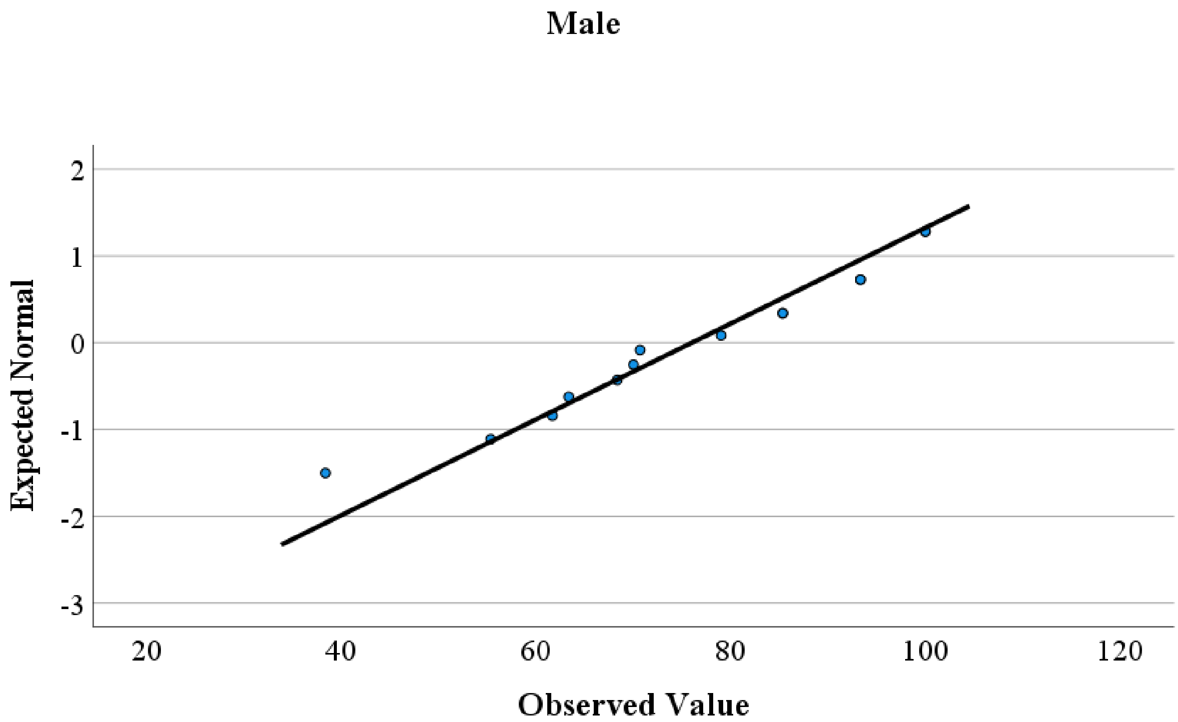


Figure 8

Q-Q Plots for Behavioral Intention Average by Gender





Appendix B - Results

Table 6

Independent Samples T-Test Results

	Practitioner		General Public						95% Confidence Interval of the Difference		
	M	*SD*	*M*	*SD*	*t*(122)	*p* (one-sided)	Mean Difference	Std. Error Difference	Lower	Upper	Cohen's *d*
BI Average	68.7	18.3	74.6	18.2	-1.8	.037	-5.91	3.28	-12.41	.59	0.323

Note. BI refers to Behavioral Intention.

Table 7

Three-Way Between-Subjects ANOVA

Predictor	Type III Sum of Squares	*df*	Mean Square	*F*	*p*	Partial η^2	Observed Power[a]
(Intercept)	135161.33	7	135161.33	281.59	<.001	.810	1.0
Gender	341.17	1	341.17	.71	.402	.011	.132
NF Experience	888.78	1	888.78	1.85	.178	.027	.268
Age	276.91	1	276.91	.58	.450	.009	.116
Gender x NF Experience	168.11	1	168.11	.35	.556	.005	.090
Gender x Age	147.32	1	147.32	.31	.581	.005	.085
Age x NF Experience	6.36	1	6.36	.01	.909	.000	.051
Gender x Age x NF Experience	286.40	1	286.40	.60	.443	.009	.119
Error	31679.89	66	479.998				

Note. R Squared = .079 (Adjusted R Squared = -.018). NF refers to neurofeedback.

[a]Computed using alpha = .05