Sensitivity to change and convergent validity of the Tinnitus Functional Index (TFI) and the Tinnitus Questionnaire (TQ) : clinical and research perspectives

Sensitivity to change and convergent validity of the Tinnitus Functional Index (TFI) and the Tinnitus Questionnaire (TQ): clinical and research perspectives

Jacquemin Laure[1,2], Mertens Griet[1,2], Van de Heyning Paul[1,2], Vanderveken Olivier[1,2], Topsakal Vedat [1,2], De Hertogh Willem [3], Michiels Sarah[1,2,3], Van Rompaey Vincent[1,2], Gilles Annick[1,2,4]

[1] University Dept. of Otorhinolaryngology and Head & Neck surgery, Antwerp University Hospital, Edegem, Belgium
[2] Dept. of Translational Neurosciences, Faculty of Medicine and Health Sciences, University of Antwerp, Wilrijk, Belgium
[3] Dept. of Rehabilitation Sciences and Physiotherapy, Faculty of Medicine and Health Sciences, University of Antwerp, Wilrijk, Belgium
[4] Dept. of Education, Health & Social Work, University College Ghent, Ghent, Belgium

laure.jacquemin@uza.be, griet.mertens@uza.be, paul.vandeheyning@uza.be, olivier.vanderveken@uza.be,

vedat.topsakal@uza.be, willem.dehertogh@uantwerpen.be, sarah.michiels@uantwerpen.be,

vincent.vanrompaey@uza.be, annick.gilles@uza.be

Corresponding author

Laure Jacquemin, Dienst NKO Wilrijkstraat 10 2650 Edegem, +32 3 821 4233, laure.jacquemin@uza.be

Conflicts of Interest and Source of Funding

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

The present research is financially supported by VLAIO (Agentschap Innoveren en Ondernemen) and a research grant from the FWO (Fonds voor Wetenschappelijk onderzoek Vlaanderen, Egmontstraat 5, 1000 Brussels) (T001916 N).

Abstract

Background

Complete alleviation of tinnitus perception is rarely achieved by current therapeutic interventions. However, adequate therapy may induce a small or large decrement of the tinnitus. The assessment of the therapeutic effect is challenging due to large interindividual variability. Multiple tinnitus questionnaires are available to describe functional effects of tinnitus pre and post-therapeutic intervention, of which the Tinnitus Functional Index (TFI) and Tinnitus Questionnaire (TQ) are two very commonly used questionnaires by clinicians and researchers.

Objective

To recommend either TFI or TQ as an outcome measure for the investigation of tinnitus treatments by investigating the consistency between the TFI and TQ, as well as comparing the responsiveness of both scales after therapeutic intervention.

Methods

Data from the validated Dutch versions of the TFI and the TQ of 100 chronic subjective tinnitus patients who underwent 6 sessions of neuromodulation at the Antwerp University Hospital were included. The patients reported their perceived effect and filled out the TFI and TQ at the pre-therapy, post-therapy and follow-up visit (i.e. ± seven weeks post-therapy). The intra-class correlation (ICC) was determined, measuring consistency between the TFI and TQ. Furthermore, the responsiveness was compared by use of mixed effects analysis. Finally, the agreement between the total scores of the questionnaires and the perceived effect was evaluated by use of their receiver operating characteristic (ROC) curve.

Results

The total scores of the TFI and TQ showed a good agreement at pre-therapy (ICC = .82) and post-therapy visit (ICC = .82). The ICC indicated a lower, but still good agreement at follow-up (ICC = .79). Furthermore, the total scores of the TFI and TQ significantly changed over time ($p_{TFI} < .05$; $p_{TQ} < .05$). In the patients who verbally reported at follow-up visit that their tinnitus complaints decreased after treatment, the percentage of patients showing a clinically significant decrease on the TFI (i.e. a decrease of ≥ 13 points) and the TQ (i.e. decrease of ≥ 12 points) was 50% and 33% respectively, however, not statically significant different ($p = .13$). Finally, the analysis of the agreement between the perceived effect and the clinically significant improvement on the questionnaires showed a kappa-value of .36 for the TFI and .29 for the TQ.

Conclusion

The TFI and TQ have high convergent validity and are both suitable as an outcome measure to quantify a change in the self-perceived tinnitus burden. The TFI is shorter and shows a slightly higher agreement with the self-reported perceived effect. Therefore, this comparison is in favor of the TFI for research purposes.

Keywords: tinnitus, tinnitus questionnaire, outcome assessment, convergent validity, responsiveness

1. Introduction

Subjective tinnitus, the sensation of a sound in the absence of an external source, is a symptom of which the prevalence is estimated between 10% and 15% in the adult population (Baguley et al., 2013). Diverse problems are associated with tinnitus, including elevated stress, anxiety, depression, sleep, concentration, and listening difficulties (Hall et al., 2018c). This phantom auditory perception can only be measured indirectly. Therefore, there are multiple tinnitus self-report questionnaires available, assessing the tinnitus severity by questioning the tinnitus-related problems (Henry, 2016; Meikle et al., 2008).

International guidelines have been developed in order to provide recommendations for the choice of self-report questionnaires in tinnitus research. The Tinnitus Questionnaire (TQ) (Hallam et al., 1988) is one of the instruments that is often recommended and used in various clinical trials (e.g. Cima et al., 2019; Fuller et al., 2017; Hesse et al., 2001; Jacquemin et al., 2018; Kleinjung et al., 2005; Landgrebe et al., 2012; Langguth et al., 2014; Langguth et al., 2007; Mazurek et al., 2006; McCombe et al., 2001; Mertens et al., 2016; Punte et al., 2011). Besides the TQ, the Tinnitus Handicap Inventory (THI) (Newman et al., 1996) is also a widely recommended tinnitus instrument. Comparative studies between the THI and TQ show high convergent validity, implying that studies using either THI or TQ are comparable (Baguley et al., 2000; Zeman et al., 2012).

The goal of the development of the TQ was to measure the most commonly reported adverse effects of tinnitus and to investigate how many underlying dimensions of complaint could be identified by conducting factor analysis (Hallam, 2008; Hallam et al., 1988). The TQ was developed in a context where cognitive therapy for tinnitus was undertaken and was standardized on a population who reported tinnitus at a neuro-otology clinic. The TQ can be used to examine relationships between different aspects of complaint and other psychological variables (Fackrell et al., 2014; Hallam, 2008), as the final 52-item TQ assesses the main psychological effects of tinnitus by covering five domains (i.e. emotional and cognitive distress, intrusiveness, auditory difficulties, sleep disturbance, and somatic complaints) (Hallam et al., 1988; Meeus et al., 2007).

The TQ and most of the other tinnitus questionnaires were developed for diagnostic purposes or classification of tinnitus severity at one point in time. Nevertheless, they are often used to asses therapeutic benefit (Landgrebe et al., 2012). In tinnitus research, therapeutic benefit is assessed in terms of relative improvement, since complete alleviation of tinnitus distress is rarely achieved by current therapeutic interventions (Van de Heyning et al., 2015). Yet, most of these questionnaires were not specifically designed to maximize responsiveness to treatment-related change (Kamalski et

al., 2010; Meikle et al., 2008). The TQ, for example, uses a three-point response scale which probably limits the responsiveness to change and thus its utility as an outcome measure (Fackrell et al., 2014).

Contrary to the TQ, the Tinnitus Functional Index (TFI), which is also a commonly used questionnaire to assess the functional effects of tinnitus (i.e. the associated complaints) (Landgrebe et al., 2012), is developed with the purpose to detect treatment-induced changes (e.g. Cima et al., 2019; Fuller et al., 2017; Landgrebe et al., 2012). As Meikle et al. (2012, 1) pointed out "Effective tinnitus treatments are urgently needed, but evaluating them is hampered by the lack of standardized measures that are validated for both intake assessment and evaluation of treatment outcomes". The development of the TFI was conducted by including a large group of tinnitus patients who were seeking clinical interventions at one of the three data collection sites. It was an iterative process that evaluated successively refined questionnaire versions. The emphasis was placed on high sensitivity to change, comprehensive coverage of the major function impacts from tinnitus and high construct validity for scaling tinnitus severity. The final TFI consists of 25 items assessing the severity and negative impact of tinnitus by covering eight domains (i.e. intrusiveness, sense of control, cognitive complaints, sleep disturbance, auditory difficulties, relaxation, quality of life and emotional distress) (Meikle et al., 2012). It has previously been observed that the auditory subscale of the TFI does not contribute to the functional impact of tinnitus (Fackrell et al., 2016; Meikle et al., 2012). Fackrell et al. (2016) suggest investigating if the auditory subscale could be calculated separately. In addition, these authors recommended not relying on the quality of life subscale of the TFI for a meaningful interpretation of generic quality of life. Furthermore, with a 10-point response scale, the TFI should be able to detect small changes in different tinnitus domains and potentially avoiding large floor and ceiling effects. Nevertheless, the responsiveness to change should be further investigated, as Fackrell et al. (2016) suggested that the degree of validity and reliability can differ across different tinnitus populations. Finally, Meikle et al. (2012) also compared the TFI with the THI. They argue that the advantages of the TFI are the ten-point response format (versus the three-level format of the THI), greater responsiveness, extensive emphasis on content validity and the eight subscales (versus three subscales of the THI).

In brief, the TFI and TQ discriminate both between levels of tinnitus distress, but only the TFI was designed to provide a responsive measure for treatment-related changes (Fackrell et al., 2014; Fackrell et al., 2016).

1.3 Research aims

The current study compares the validated Dutch versions of the TQ (a questionnaire recommended by international guidelines) and the TFI (a questionnaire specifically developed to be sensitive to

change). The hypothesis is that the TFI is a more sensitive instrument when assessing changes as a result of an intervention. Hence, the goal is to recommend the most appropriate outcome measure for the investigation of tinnitus treatments in research. To our knowledge, the comparison between the TFI and TQ as a therapy assessment tool has never been conducted before.

2. Material and methods

2.1 Study design

The current study is a retrospective study conducted at the Antwerp University Hospital. The validated Dutch versions of the TFI and TQ were filled out for research purposes at pre-therapy (i.e. before the start of the therapy), at post-therapy (i.e. three weeks after the start of the therapy) and at follow-up (i.e. ± seven weeks after ending the therapy) using a touch-screen desktop. The therapy consisted of six sessions of High-Definition transcranial Direct Current Stimulation (HD-tDCS) of the right dorsolateral prefrontal cortex, with two sessions of 20 minutes weekly. The scoring of the questionnaires was computerized. Besides collecting the questionnaire data, the perceived effect was also questioned at the follow-up visit.

2.2 Subjects

A total of 100 chronic (≥ 6 months), subjective, non-pulsatile tinnitus patients who underwent HD-tDCS were included in the present study. They were screened with regard to the safety of applying HD-tDCS (Villamar et al., 2013) and excluded if they had a middle ear pathology or were undergoing another tinnitus treatment. The demographic details of the participants are presented in Table 1. The total scores of the TFI and TQ prior therapy are shown in Figure 1. Table 2 provides the mean and standard deviation (SD) of the TFI and TQ total scores and subscales at each time point.

[Figure 1 near here]

Gender	female: n = 17, male: n = 83
Age at pre-therapy	$\bar{x}$ = 51 years [19 – 80 years]
PTA	$PTA_{low\ right}$ $\bar{x}$ = 12 dB HL [-3 – 88 dB HL] $PTA_{high\ right}$ $\bar{x}$ = 18 dB HL [-5 – 88 dB HL] $PTA_{low\ left}$ $\bar{x}$ = 12 dB HL [-3 – 48 dB HL] $PTA_{high\ left}$ $\bar{x}$ = 19 dB HL [-2 – 63 dB HL]
Duration tinnitus	$\bar{x}$ = 6.2 years [0.6 – 34 years]
Side tinnitus	bilateral: n = 59, central: n = 15, left: n = 14, right: n = 12
Type tinnitus	pure tone: n = 68, noise: n = 21, polyphonic: n = 11
Onset of tinnitus	spontaneous: n = 27, ear pathology: n = 46, non-otologic: n = 4, psychological: n = 4, unknown: n = 19

Table 1: Demographic details of the participants. Polyphonic tinnitus was defined as 'a mixture of different sounds'. The onset of tinnitus was determined after an interview with the participants. They

reported that the tinnitus suddenly appeared in the absence of a specific event (spontaneous) or was noticed after a decrease of hearing (ear pathology), physical problem outside the ear (non-otologic) or a psychological event (psychological). The onset was unclear in 19 participants (unknown). Other causes were questioned but not applicable to these participants (PTA, pure tone average; PTA_{low}, 0.5-1-2 kHz; PTA_{high}, 1-2-4 kHz)

	Pre-therapy	Post-therapy	Follow-up
TFI total score	**$\bar{x}$ = 47 (SD: 20)**	**$\bar{x}$ = 45 (SD: 20)**	**$\bar{x}$ = 42 (SD: 20)**
TFI Intrusiveness	$\bar{x}$ = 62 (SD: 22)	$\bar{x}$ = 59 (SD: 22)	$\bar{x}$ = 56 (SD: 23)
TFI control	$\bar{x}$ = 59 (SD: 20)	$\bar{x}$ = 57 (SD: 22)	$\bar{x}$ = 51 (SD: 22)
TFI cognitive	$\bar{x}$ = 42 (SD: 23)	$\bar{x}$ = 40 (SD: 23)	$\bar{x}$ = 36 (SD: 22)
TFI sleep	$\bar{x}$ = 42 (SD: 31)	$\bar{x}$ = 43 (SD: 30)	$\bar{x}$ = 41 (SD: 31)
TFI auditory	$\bar{x}$ = 37 (SD: 26)	$\bar{x}$ = 39 (SD: 25)	$\bar{x}$ = 36 (SD: 24)
TFI relaxation	$\bar{x}$ = 59 (SD: 26)	$\bar{x}$ = 57 (SD: 25)	$\bar{x}$ = 54 (SD: 27)
TFI quality of life	$\bar{x}$ = 37 (SD: 26)	$\bar{x}$ = 34 (SD: 23)	$\bar{x}$ = 33 (SD: 24)
TFI emotional	$\bar{x}$ = 42 (SD: 26)	$\bar{x}$ = 37 (SD: 24)	$\bar{x}$ = 36 (SD: 25)
TQ total score	**$\bar{x}$ = 36 (SD: 16)**	**$\bar{x}$ = 36 (SD: 15)**	**$\bar{x}$ = 34 (SD: 14)**
TQ distress	$\bar{x}$ = 17 (SD: 9)	$\bar{x}$ = 17 (SD: 8)	$\bar{x}$ = 16 (SD: 8)
TQ auditory	$\bar{x}$ = 4 (SD: 3)	$\bar{x}$ = 4 (SD: 3)	$\bar{x}$ = 4 (SD: 3)
TQ intrusiveness	$\bar{x}$ = 10 (SD: 3)	$\bar{x}$ = 10 (SD: 3)	$\bar{x}$ = 9 (SD: 4)
TQ sleep	$\bar{x}$ = 3 (SD: 3)	$\bar{x}$ = 3 (SD: 3)	$\bar{x}$ = 3 (SD: 3)
TQ somatic	$\bar{x}$ = 2 (SD: 2)	$\bar{x}$ = 2 (SD: 2)	$\bar{x}$ = 2 (SD: 2)

Table 2: Mean and standard deviation (SD) of the TFI and TQ total scores and the subscales at each time point *(Tinnitus Functional Index, TFI; Tinnitus Questionnaire, TQ).*

2.3 Measurements

The 52 items of the TQ assesses the main psychological effects of tinnitus by covering five domains (i.e. emotional and cognitive distress, intrusiveness, auditory difficulties, sleep disturbance, and somatic complaints) (Hallam et al., 1988; Meeus et al., 2007). The patient must indicate the level of agreement by answering; not true (score 0), partly true (score 1) or true (score 2). As 12 out of the 52 items are excluded from the scoring (items 6, 23, 24, 29, 30, 32, 40, 42, 45, 46, 49, and 52) and two items must be scored twice (items 5 and 20), the total score ranges from 0-84, with higher scores denoting higher levels of distress. An improvement of 12 points or more has been considered as a clinically significant improvement (Hall et al., 2018a), meaning an improvement noticeable for the patient. The psychometric characters of the Dutch version of the TQ are similar to the original English questionnaire, with a high Cronbach's alpha value of 0.95 (Meeus et al., 2007).

The TFI is a 25-items questionnaire scoring the severity and negative impact of tinnitus by covering eight domains (i.e. intrusiveness, sense of control, cognitive complaints, sleep disturbance, auditory difficulties, relaxation, quality of life and emotional distress). The patient must answer each item on a Likert scale from 0 to 10. The TFI discriminates between levels of tinnitus distress, with the total

score ranging from 0-100 and higher scores indicating higher levels of disturbance. An improvement of 13 points or more on the TFI has been indicated as a clinically significant improvement (Meikle et al., 2012). The internal consistency of the Dutch version of the TFI is very high with a Cronbach's alpha value of 0.96, which is similar to the original (Rabau et al., 2014).

The perceived effect (i.e. self-report) was also determined at follow-up visit by asking an open-ended question: "To which extent did you perceive a change in your tinnitus perception during and after the HD-tDCS treatment?".

2.4 Statistical analysis

The objective of the present study was to recommend either the TFI or TQ as the most appropriate outcome measure for the investigation of tinnitus treatments in research.

Data were analyzed using SPSS statistical software version 25 (SPSS Inc., Chicago, IL, USA).

The normality of the dataset was evaluated by performing the Shapiro-Wilk test. In addition, the normality was determined by visualizing the data in histograms. The normality of the data was confirmed ($p > .05$). Consequently, parametric tests with mean and SD of the variables are reported. The TQ data were missing for five patients at the pre-therapy visit and for one patient at the post-therapy and follow-up visit.

2.4.1 Convergent validity

The first research question was: "Do the TFI and TQ measure the same construct?". This was investigated by assessing the convergent validity between the total scores of the TFI and TQ, and the convergent validity between the similar subscales of the TFI and TQ (i.e. intrusiveness, auditory and sleep) at the pre-therapy, post-therapy and follow-up appointment**.** The *convergent validity* is the relationship between two questionnaires aiming to measure the same hypothetical construct (e.g. tinnitus distress). It was expected that total scores of the TFI and TQ, and the subscales of the TFI and TQ with the same label would show a high convergent validity. In addition, the discriminant validity of the other subscales of the TFI and TQ, which were expected to measure different constructs, was evaluated. The *discriminant validity* is the lack of correlation between measurements of unrelated variables.

A two-way-mixed intraclass correlation (ICC) was determined, with an ICC of -1 indicating perfect disagreement, 0 indicating completely random agreement and 1 indicating perfect agreement according to the guidelines of Koo et al. (2016). More specifically, values less than 0.50 indicate poor agreement, values between 0.50 and 0.75 indicate moderate agreement, values between 0.75 and

0.90 indicate good agreement, and values greater than 0.90 indicate excellent agreement (Koo et al., 2016).

2.4.2 Responsiveness

The second research question was: “Is the responsiveness of the TFI higher compared to the TQ?”. The *responsiveness* of a questionnaire indicates the ability to accurately detect change over time (Mokkink et al., 2012). This latter is often determined by comparison with the effect perceived by the patient. Multiple analyses were required to provide an answer to this research question.

First, the change in TFI or TQ after treatment was investigated using mixed-effects analysis to account for the non-independence between the observations taken from the same individual. The total scores and subscales of the TFI and TQ were entered as a dependent variable and a random intercept for individuals was included. Post-hoc comparisons were performed with Bonferroni correction. The significance level was set at $p < .05$. In addition, the participants were divided into responders and non-responders, according to their perceived effect at follow-up. T-tests were conducted to investigate if the change in TFI and TQ total scores and subscales (i.e. from pre-therapy to follow-up) differed between these two groups.

Subsequently, the Cohen’s kappa was calculated to provide a measure of agreement between the perceived effect reported by the patients and the clinically significant change in either the total scores of the TFI or TQ (Watson et al., 2010). This value was interpreted according to the guidelines of Landis and Koch (1977). More specifically, values less than 0 indicate poor agreement, values between 0 and 0.20 indicate slight agreement, values between 0.21 and 0.40 indicate fair agreement, values between 0.41 and 0.60 indicate moderate agreement, values between 0.61 and 0.80 indicate substantial agreement and values higher than 0.80 indicate almost perfect agreement.

Finally, the area under the receiver operating characteristics (ROC) curve (AUC) was calculated to evaluate the ability of the following variables to distinguish between responders and non-responders: ‘the change in the total scores of either TFI or TQ from pre-therapy to follow-up visit’ and ‘the total score of either TFI or TQ at follow-up visit’. A separate ROC analysis was conducted for each variable. According to the recommendations of Terwee et al. (2007), the AUC should be at least 0.70 to be adequate. In addition, the McNemar’s test evaluated the difference between the total scores of the TFI and TQ to distinguish between these two groups.

2.5 Ethics committee approval

The Committee for Medical Ethics of the Antwerp University Hospital approved the study (file number: 16/41/415). All participants provided written informed consent.

3. Results

3.1 Convergent validity

The intra-class correlation (ICC), measuring consistency between the total scores of the TFI and TQ, showed a good agreement at pre-therapy (ICC = .82) and post-therapy visit (ICC = .82). The ICC indicated a lower, but still good agreement at follow-up (ICC = .79) (Table 3).

TFI and TQ total scores	ICC	95% CI
Pre-therapy	.82	[.74 ; .87]
Post-therapy	.82	[.74 ; .87]
Follow-up	.79	[.70 ; .85]

	TQ intrusive	TQ auditory	TQ sleep	TQ distress	TQ somatic
TFI intrusive	.22 [.02 ; .40]	.11 [-.09 ; .30]	.07 [-.13 ; .27]	.44 [.26 ; .58]	.07 [-.13 ; .27]
TFI control	.22 [.02 ; .40]	.12 [-.09 ; .31]	.08 [-.12 ; .28]	.48 [.31 ; .62]	.07 [-.13 ; .27]
TFI cognitive	.19 [-.01 ; .38]	.16 [-.05 ; .35]	.06 [-.15 ; .25]	.40 [.21 ; .55]	.07 [-.14 ; .26]
TFI sleep	.10 [-.10 ; .30]	.06 [-.15 ; .26]	.14 [-.07 ; .33]	.28 [.09 ; .46]	.05 [-.15 ; .25]
TFI auditory	.13 [-.07 ; .32]	.18 [-.03 ; .36]	.01 [-.20 ; .21]	.20 [.00 ; .38]	.03 [-.17 ; .23]
TFI relaxation	.19 [-.02 ; .37]	.08 [-.13 ; .27]	.08 [-.13 ; .27]	.39 [.21 ; .55]	.05 [-.15 ; .25]
TFI QOL	.19 [-.01 ; .38]	.15 [-.05 ; .34]	.07 [-.14 ; .26]	.43 [.25 ; .58]	.07 [-.14 ; .26]
TFI emotional	.20 [.00 ; .39]	.12 [-.08 ; .32]	.07 [-.13 ; .27]	.47 [.30 ; .62]	.06 [-.14 ; .26]

Table 3: Consistency analysis between the TFI and TQ total scores at pre-therapy, post-therapy and follow-up visit (upper). Consistency analysis between the TFI and TQ subscales at pre-therapy (below). *(CI, confidence interval, ICC, intra-class correlation, Tinnitus Functional Index, TFI; Tinnitus Questionnaire, TQ).*

The analysis of the convergent validity of the similar subscales of the TFI and TQ (i.e. intrusiveness, auditory and sleep) showed a poor agreement at the pre-therapy visit ($ICC_{intrusiveness}$ = .22; $ICC_{auditory}$ = .18; ICC_{sleep} = .14) (Table 3). The discriminant validity of the other subscales of the TFI and TQ, which were expected to measure different constructs, indicated also poor agreements (ICC < .50) (Table 3). The observed agreement between the TFI and TQ subscales was consistent over time.

3.2 Responsiveness

[Figure 2 near here]

The total scores of the TFI and TQ significantly changed over time ($p_{TFI} < .05$; $p_{TQ} < .05$) (Figure 2). Post-hoc comparisons showed a significant improvement on the TFI total score between the pre-therapy and follow-up visit ($p < .05$). The improvement on the TQ total score between these two visits was not significant ($p > .05$). Moreover, the change in the TFI and TQ total score (i.e. from pre-therapy to follow-up) differed significantly between responders and non-responders ($p < .01$) (Table

4). This subdivision was made in terms of perceived effect. The change in the TFI subscales scores showed that four out of eight TFI subscales improved significantly over time (i.e. from pre-therapy to post-therapy to follow-up), namely intrusiveness ($p_{time} < .05$; $p_{pre\text{-}fu} < .05$), sense of control ($p_{time} < .01$; $p_{pre\text{-}fu} < .01$; $p_{post\text{-}fu} < .01$), cognitive complaints ($p_{time} < .05$; $p_{pre\text{-}fu} < .05$) and emotional distress ($p_{time} < .05$; $p_{pre\text{-}post} < .05$; $p_{pre\text{-}fu} < .05$). Moreover, the responders and non-responders differed significantly from each other in terms of the change in TFI subscales (i.e. from pre-therapy to follow-up) ($p < .01$ for all subscales, except for sleep and auditory $p < .05$) (Table 4). The intrusiveness subscale was the only TQ subscale showing a significant improvement over time ($p_{time} < .05$; $p_{pre\text{-}fu} < .05$; $p_{post\text{-}fu} < .05$). Similarly to the TFI subscales, the TQ subscales differed significantly between responders and non-responders ($p < .01$ for all subscales, except for somatic $p < .05$) (Table 4).

	Mean change Non-responder	Mean change responder	P-value
TFI total score	.18	14.55	<.001
TFI Intrusiveness	1.05	16.33	<.001
TFI control	3.86	18.00	.001
TFI cognitive	2.00	15.44	.002
TFI sleep	-2.67	11.45	.011
TFI auditory	-1.62	7.45	.048
TFI relaxation	-.05	15.44	<.001
TFI quality of life	-1.69	15.00	.001
TFI emotional	1.24	17.11	.002
TQ total score	-.75	9.57	.001
TQ distress	-.02	4.30	.002
TQ auditory	-.17	1.23	.006
TQ intrusiveness	-.17	2.70	<.001
TQ sleep	-.23	.93	.005
TQ somatic	-.16	.40	.043

Table 4: Change in TFI and TQ total scores and subscales (i.e. from pre-therapy to follow-up) for non-responders (n = 70) and responders (n = 30). The p-values correspond to the independent sample t-tests comparing the change between responders and non-responders.

The analysis of the agreement between the subjectively perceived effect, reported by the patients, and the clinically significant improvement on the questionnaires, showed a kappa-value of .36 for the TFI and .29 for the TQ, meaning that both questionnaires showed a fair agreement with the perceived effect.

The ROC curve showed an AUC of .734 (95% CI, .63 to .84) for the change on the TFI and .730 (95% CI, .62 to .84) for the change on the TQ. The AUC for the absolute total score of the TFI and TQ at follow-up visit was .676 (95% CI, .56 to .80) and .632 (95% CI, .50 to .76) respectively. In other words, the change in the TFI or TQ had the ability to adequately distinguish responders from non-responders.

Among the responders (n = 30), 80% and 77% showed an improvement on the total scores of the TFI and TQ respectively (Figure 3). However, this difference between TFI and TQ was not significant ($p > .05$). Furthermore, the improvement on the TFI and TQ was clinically significant for 50% and 33% of the responders respectively, but this difference between TFI and TQ was not significant ($p > .05$).

[Figure 3 near here]

4. Discussion

The current study investigated the superiority of the Dutch TFI over the Dutch TQ in terms of responsiveness to treatment-related changes. First, the convergent validity of these questionnaires was determined to investigate if they measure the same construct. Second, the responsiveness to the therapeutic effects of an HD-tDCS treatment was compared between these questionnaires.

The convergent validity of the TQ and TFI total scores was consistently high over time, indicating that the total scores reflect a similar general construct for tinnitus-related problems. Regarding the subscales of both questionnaires, the convergent validity between similar subscales of the TFI and TQ was much lower compared to the total scores. As expected, the other subscales did also showed poor agreements. Even though, there was a large variation in agreement scores between the subscales. More specifically, one of the highest agreements was found between the TQ emotional and cognitive distress subscale and the TFI intrusiveness subscale. These findings suggest that, even though the TFI and the TQ have similar labels for some subscales, they are measuring different underlying aspects of intrusiveness, sleep disturbance, and auditory difficulties. In other words, studies using either the total score of the TFI or the TQ can be compared, but the scores of the subscales of these questionnaires can differ. This can be attributed to a different amount of items assessing particular functional effects of tinnitus. For example, the TQ does not assess the interference with participation, while the TFI does not assess the beliefs about the tinnitus or the somatic complaints. Hence, researchers and clinicians may prefer either the TFI or TQ depending on the purpose.

Tinnitus intrusiveness is an important concept in the evaluation of tinnitus treatment (Hall et al., 2018b), but each questionnaire measures intrusiveness differently. This suggests that the construct of intrusiveness is complex and that it cannot be measured unambiguously. Therefore, there is a need for consensus on the meaning of the term ‘intrusiveness’. Intrusiveness is covered by the TFI and TQ, as both questionnaires comprise a subscale for this concept. However, the TQ provides more measurement points for intrusiveness than the TFI, which results in a slightly higher weight of this concept to the TQ total score. Nevertheless, the importance of measuring tinnitus intrusiveness was

confirmed in the current study, as the intrusiveness subscales for both the TQ and TFI showed improvements over time after treatment.

The second research aim was to evaluate the change over time for the TFI and TQ after neuromodulation. Both questionnaires significantly changed over time, however, there was a larger change on the TFI. Moreover, the TFI showed significant post-hoc comparisons, with the TFI total score improving significantly between pre-therapy and follow-up visit. The evaluation of the change in subscales scores, on the other hand, showed that four out of eight TFI subscales (intrusiveness, cognitive, control, emotional) and the TQ intrusiveness subscale improved significantly. But, an uncontrolled factor is a possibility that the changes over time could be related to the measurement error associated with each questionnaire. The result of the post-hoc comparisons stresses the importance of also assessing the therapy outcome during a follow-up visit, as the changes were most salient at that point. At the follow-up visit, there is also less influence of the act of undergoing therapy. Finally, the responders and non-responders differed significantly from each other in terms of the change in TFI and TQ scores. When focusing on the responders, the percentage of patients showing a clinically significant improvement on the TFI was higher compared to the TQ. However, there was no significant difference between the TFI and TQ in the number of responders showing a clinically significant improvement. Also, the ROC curve showed that both questionnaires were adequate in assessing a change in tinnitus perception. However, this study should be repeated with different tinnitus populations, as suggested by Fackrell et al. (2016). The use of 'the change in the questionnaire's total score' as outcome measurement was supported by the analyses, as the AUC was higher for 'the change in TFI/TQ total score' compared to 'the absolute TFI/TQ total score at follow-up'. In sum, the results show a trend for the TFI being more responsive, but this was not significantly different from the TQ. However, together with the fact that the TFI is shorter, we recommend using the change in TFI total score as an outcome measure for research purposes.

Important insights for future studies can be provided by evaluating the perceived effect by use of goal attainment scaling (Kiresuk et al., 1968; Kiresuk et al., 1982). It evaluates treatments in the framework of goals, which have been set before treatment together with the patient. After the termination of the treatment, the patient is asked to indicate the extent to which the goals have been achieved. The attainment of these goals can be transformed into a standardized score. This suggestion is especially important for these kinds of studies with small therapeutic effects. Neuromodulation may not result in large improvements for every patient and no consensus has been reached on what effect it has on tinnitus. Moreover, the anchor point for a clinically significant change should be further investigated for both the TFI and TQ, as this research question could not be addressed due to the small therapeutic effect in the current study.

In conclusion, the current study indicated that the TFI and TQ have high convergent validity and are both suitable as outcome measures to quantify a change in the self-perceived tinnitus burden. For research aiming to investigate the effect of therapeutic interventions, the current study supports the higher sensitivity to change of the TFI, which confirms our expectations since the TFI is designed to detect treatment-related changes by using a Likert scale for responses. Together with the fact that the TFI is shorter, this comparison is in favor of the TFI for research purposes. Replication of this study with other tinnitus treatments and other subtypes of tinnitus patients, such as somatic tinnitus, may contribute to the understanding of the full picture. Future research may focus on further comparing the TFI with other questionnaires, such as the THI, as these are both short, widely used questionnaires.

Acknowledgments

We gratefully acknowledge the statistical support of Kristien Wouters.

Funding

The present research is financially supported by VLAIO (Agentschap Innoveren en Ondernemen) and a research grant from the FWO (Fonds voor Wetenschappelijk onderzoek Vlaanderen, Egmontstraat 5, 1000 Brussels) (T001916 N).

References

Baguley, D., McFerran, D., Hall, D. 2013. Tinnitus. Lancet (London, England) 382, 1600-7.

Baguley, D.M., Humphriss, R.L., Hodgson, C.A. 2000. Convergent validity of the tinnitus handicap inventory and the tinnitus questionnaire. The Journal of laryngology and otology 114, 840-3.

Cima, R.F.F., Mazurek, B., Haider, H., Kikidis, D., Lapira, A., Noreña, A., Hoare, D.J. 2019. A multidisciplinary European guideline for tinnitus: diagnostics, assessment, and treatment. HNO 67, 10-42.

Fackrell, K., Hall, D.A., Barry, J.G., Hoare, D.J. 2014. Tools for tinnitus measurement: development and validity of questionnaires to assess handicap and treatment effects. In: Signorelli, F., Turjman, F., (Eds.), Tinnitus: Causes, Treatment and Short & Long-Term Health Effects. . Nova Science Publishers Inc, New York. pp. 13-60.

Fackrell, K., Hall, D.A., Barry, J.G., Hoare, D.J. 2016. Psychometric properties of the Tinnitus Functional Index (TFI): Assessment in a UK research volunteer population. Hearing research 335, 220-235.

Fuller, T.E., Haider, H.F., Kikidis, D., Lapira, A., Mazurek, B., Norena, A., Rabau, S., Lardinois, R., Cederroth, C.R., Edvall, N.K., Brueggemann, P.G., Rosing, S.N., Kapandais, A., Lungaard, D., Hoare, D.J., Cima, R.F.F. 2017. Different Teams, Same Conclusions? A Systematic Review of Existing Clinical Guidelines for the Assessment and Treatment of Tinnitus in Adults. Frontiers in Psychology 8.

Hall, D.A., Mehta, R.L., Argstatter, H. 2018a. Interpreting the Tinnitus Questionnaire (German version): what individual differences are clinically important? International journal of audiology 57, 553-557.

Hall, D.A., Smith, H., Hibbert, A., Colley, V., Haider, H.F., Horobin, A., Londero, A., Mazurek, B., Thacker, B., Fackrell, K. 2018b. The COMiT'ID Study: Developing Core Outcome Domains Sets for Clinical Trials of Sound-, Psychology-, and Pharmacology-Based Interventions for Chronic Subjective Tinnitus in Adults. Trends in hearing 22, 2331216518814384.

Hall, D.A., Fackrell, K., Li, A.B., Thavayogan, R., Smith, S., Kennedy, V., Tinoco, C., Rodrigues, E.D., Campelo, P., Martins, T.D., Lourenço, V.M., Ribeiro, D., Haider, H.F. 2018c. A narrative synthesis of research evidence for tinnitus-related complaints as reported by patients and their significant others. Health and quality of life outcomes 16, 61.

Hallam, R.S. 2008. TQ Manual of the Tinnitus Questionnaire: Revised and updated Polpresa Press, London.

Hallam, R.S., Jakes, S.C., Hinchcliffe, R. 1988. Cognitive variables in tinnitus annoyance. The British journal of clinical psychology 27 (Pt 3), 213-22.

Henry, J.A. 2016. "Measurement" of Tinnitus. Otology & neurotology : official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology 37, e276-85.

Hesse, G., Rienhoff, N.K., Nelting, M., Laubert, A. 2001. Chronic complex tinnitus: therapeutic results of inpatient treatment in a tinnitus clinic. Laryngorhinootologie 80, 503-508.

Jacquemin, L., Shekhawat, G.S., Van de Heyning, P., Mertens, G., Fransen, E., Van Rompaey, V., Topsakal, V., Moyaert, J., Beyers, J., Gilles, A. 2018. Effects of Electrical Stimulation in Tinnitus Patients: Conventional Versus High-Definition tDCS. Neurorehabilitation and neural repair 32, 714-723.

Kamalski, D.M., Hoekstra, C.E., van Zanten, B.G., Grolman, W., Rovers, M.M. 2010. Measuring disease-specific health-related quality of life to evaluate treatment outcomes in tinnitus patients: a systematic review. Otolaryngol Head Neck Surg 143, 181-5.

Kiresuk, T.J., Sherman, R.E. 1968. Goal attainment scaling: A general method for evaluating comprehensive community mental health programs. Community Mental Health Journal 4, 443-453.

Kiresuk, T.J., Lund, S.H., Larsen, N.E. 1982. Measurement of goal attainment in clinical and health care programs. Drug intelligence & clinical pharmacy 16, 145-53.

Kleinjung, T., Eichhammer, P., Langguth, B., Jacob, P., Marienhagen, J., Hajak, G., Wolf, S.R., Strutz, J. 2005. Long-Term Effects of Repetitive Transcranial Magnetic Stimulation (rTMS) in Patients with Chronic Tinnitus. Otolaryngology–Head and Neck Surgery 132, 566-569.

Koo, T.K., Li, M.Y. 2016. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. Journal of Chiropractic Medicine 15, 155-63.

Landgrebe, M., Azevedo, A., Baguley, D., Bauer, C., Cacace, A., Coelho, C., Dornhoffer, J., Figueiredo, R., Flor, H., Hajak, G., van de Heyning, P., Hiller, W., Khedr, E., Kleinjung, T., Koller, M., Lainez, J.M., Londero, A., Martin, W.H., Mennemeier, M., Piccirillo, J., De Ridder, D., Rupprecht, R., Searchfield, G., Vanneste, S., Zeman, F., Langguth, B. 2012. Methodological aspects of clinical trials in tinnitus: A proposal for an international standard. Journal of psychosomatic research 73, 112-121.

Landis, J.R., Koch, G.G. 1977. The measurement of observer agreement for categorical data. Biometrics 33, 159-74.

Langguth, B., Landgrebe, M., Frank, E., Schecklmann, M., Sand, P.G., Vielsmeier, V., Hajak, G., Kleinjung, T. 2014. Efficacy of different protocols of transcranial magnetic stimulation for the treatment of tinnitus: Pooled analysis of two randomized controlled studies. The World Journal of Biological Psychiatry 15, 276-285.

Langguth, B., Goodey, R., Azevedo, A., Bjorne, A., Cacace, A., Crocetti, A., Del Bo, L., De Ridder, D., Diges, I., Elbert, T., Flor, H., Herraiz, C., Ganz Sanchez, T., Eichhammer, P., Figueiredo, R., Hajak, G., Kleinjung, T., Landgrebe, M., Londero, A., Lainez, M.J.A., Mazzoli, M., Meikle, M.B., Melcher, J., Rauschecker, J.P., Sand, P.G., Struve, M., Van de Heyning, P., Van Dijk, P., Vergara, R. 2007. Consensus for tinnitus patient assessment and treatment outcome measurement: Tinnitus Research Initiative meeting, Regensburg, July 2006. In: Langguth, B., Hajak, G., Kleinjung, T., Cacace, A., Møller, A.R., (Eds.), Progress in brain research, Vol. 166. Elsevier. pp. 525-536.

Mazurek, B., Fischer, F., Haupt, H., Georgiewa, P., Reisshauer, A., Klapp, B.F. 2006. A Modified Version of Tinnitus Retraining Therapy: Observing Long-Term Outcome and Predictors. Audiology and Neurotology 11, 276-286.

McCombe, A., Baguley, D., Coles, R., McKenna, L., McKinney, C., Windle-Taylor, P. 2001. Guidelines for the grading of tinnitus severity: the results of a working group commissioned by the British Association of Otolaryngologists, Head and Neck Surgeons, 1999. Clinical Otolaryngology & Allied Sciences 26, 388-393.

Meeus, O., Blaivie, C., Van de Heyning, P. 2007. Validation of the Dutch and the French version of the Tinnitus Questionnaire. B-ent 3 Suppl 7, 11-7.

Meikle, M.B., Stewart, B.J., Griest, S.E., Henry, J.A. 2008. Tinnitus outcomes assessment. Trends in amplification 12, 223-35.

Meikle, M.B., Henry, J.A., Griest, S.E., Stewart, B.J., Abrams, H.B., McArdle, R., Myers, P.J., Newman, C.W., Sandridge, S., Turk, D.C., Folmer, R.L., Frederick, E.J., House, J.W., Jacobson, G.P., Kinney, S.E., Martin, W.H., Nagler, S.M., Reich, G.E., Searchfield, G., Sweetow, R., Vernon, J.A. 2012. The tinnitus functional index: development of a new clinical measure for chronic, intrusive tinnitus. Ear and hearing 33, 153-76.

Mertens, G., De Bodt, M., Van de Heyning, P. 2016. Cochlear implantation as a long-term treatment for ipsilateral incapacitating tinnitus in subjects with unilateral hearing loss up to 10 years. Hearing research 331, 1-6.

Mokkink, L.B., Terwee, C.B., Patrick, D.L., Alonso, J., Stratford, P.W., Knol, D.L., Bouter, L.M., de Vet, H.C.W. 2012. The COSMIN checklist manual. [Online] [Online]. Available by Amsterdam: VU University Medical Centre http://fac.ksu.edu.sa/sites/default/files/cosmin_checklist_manual_v9.pdf.

Newman, C.W., Jacobson, G.P., Spitzer, J.B. 1996. Development of the Tinnitus Handicap Inventory. Archives of otolaryngology--head & neck surgery 122, 143-8.

Punte, A.K., Vermeire, K., Hofkens, A., De Bodt, M., De Ridder, D., Van de Heyning, P. 2011. Cochlear implantation as a durable tinnitus treatment in single-sided deafness. Cochlear implants international 12, S26-S29.

Rabau, S., Wouters, K., Van de Heyning, P. 2014. Validation and translation of the Dutch tinnitus functional index. B-ent 10, 251-8.

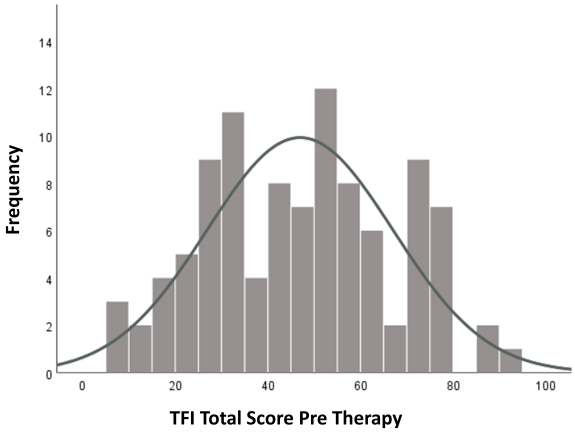
Terwee, C.B., Bot, S.D., de Boer, M.R., van der Windt, D.A., Knol, D.L., Dekker, J., Bouter, L.M., de Vet, H.C. 2007. Quality criteria were proposed for measurement properties of health status questionnaires. Journal of clinical epidemiology 60, 34-42.

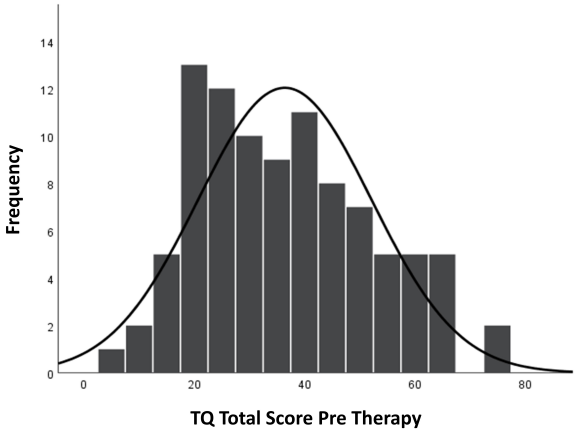
Van de Heyning, P., Gilles, A., Rabau, S., Van Rompaey, V. 2015. Subjective tinnitus assessment and treatment in clinical practice: the necessity of personalized medicine. Current opinion in otolaryngology & head and neck surgery 23, 369-75.

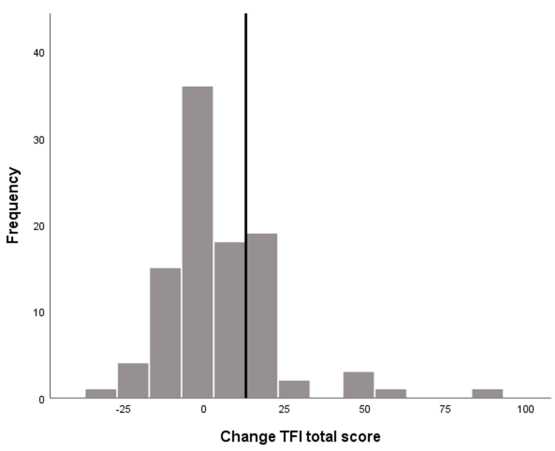
Villamar, M.F., Volz, M.S., Bikson, M., Datta, A., Dasilva, A.F., Fregni, F. 2013. Technique and considerations in the use of 4x1 ring high-definition transcranial direct current stimulation (HD-tDCS). Journal of visualized experiments : JoVE, e50309.

Watson, P.F., Petrie, A. 2010. Method agreement analysis: A review of correct methodology. Theriogenology 73, 1167-1179.

Zeman, F., Koller, M., Schecklmann, M., Langguth, B., Landgrebe, M. 2012. Tinnitus assessment by means of standardized self-report questionnaires: psychometric properties of the Tinnitus Questionnaire (TQ), the Tinnitus Handicap Inventory (THI), and their short versions in an international and multi-lingual sample. Health and quality of life outcomes 10, 128.







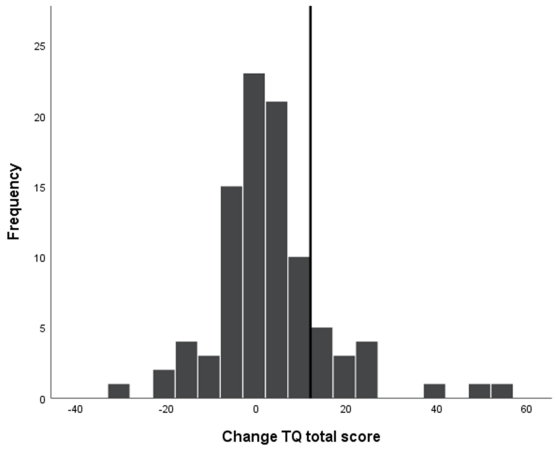
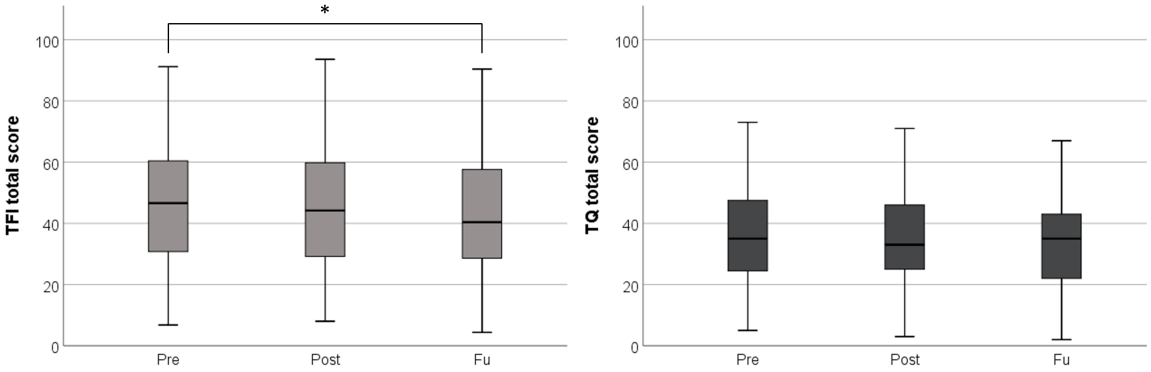


Figure 1: Histograms of the total scores of the TFI and TQ at pre-therapy visit, with the mean TFI score of 47.0 (SD = 20.1) (n = 100) and the mean TQ score of 36 (SD = 16) (n = 95). *(Tinnitus Functional Index, TFI; Tinnitus Questionnaire, TQ).*

Figure 2: Boxplots (min, Q1, median, Q3, max) representing the total scores of the TFI and TQ at pre-therapy (i.e. before the start of the therapy), at post-therapy (i.e. three weeks after the start of the therapy) and at follow-up (i.e. ± seven weeks after ending the therapy). Significant changes ($p < .05$) are indicated with an asterisk (*). *(Tinnitus Functional Index, TFI; Tinnitus Questionnaire, TQ).*

Figure 3: Histograms of the change in the TFI and TQ total score between pre-therapy and follow-up visit for the responders (n = 30) *(Tinnitus Functional Index, TFI; Tinnitus Questionnaire, TQ).*



Highlights

- High convergent validity between TFI and TQ
- TFI and TQ are suitable as outcome measures quantifying treatment-related changes
- The TFI shows a slightly higher agreement with the self-reported perceived effect.
- This comparison is in favor of the TFI for research purposes.