Received: 14 November 2020 | Revised: 3 April 2021 | Accepted: 17 May 2021

DOI: 10.1111/psyp.13874

REVIEW

PSYCHOPHYSIOLOGY SPR WILEY

EEG-neurofeedback and executive function enhancement in healthy adults: A systematic review

Giada Viviani[1] | **Antonino Vallesi**[1,2]

[1]Department of Neuroscience and Padova Neuroscience Center, University of Padova, Padova, Italy

[2]IRCCS San Camillo Hospital, Venice, Italy

Correspondence
Antonino Vallesi, Department of Neuroscience, University of Padova, via Giustiniani 2, 35128, Padova, Italy.
Email: antonino.vallesi@unipd.it

Funding information
Ministero dell'Istruzione, dell'Università e della Ricerca, Grant/Award Number: "Department of excellence 2018-2022" initiative awarded to the Department of Neuroscience – University of Padua.

Abstract

Electroencephalographic (EEG)-neurofeedback training (NFT) is a promising technique that supports individuals in learning to modulate their brain activity to obtain cognitive and behavioral improvements. EEG-NFT is gaining increasing attention for its potential "peak performance" applications on healthy individuals. However, evidence for clear cognitive performance enhancements with healthy adults is still lacking. In particular, whether EEG-NFT represents an effective technique for enhancing healthy adults' executive functions is still controversial. Therefore, the main objective of this systematic review is to assess whether the existing EEG-NFT studies targeting executive functions have provided reliable evidence for NFT effectiveness. To this end, we conducted a qualitative analysis of the literature since the limited number of retrieved studies did not allow us meta-analytical comparisons. Moreover, a second aim was to identify optimal frequencies as NFT targets for specifically improving executive functions. Overall, our systematic review provides promising evidence for NFT effectiveness in boosting healthy adults' executive functions. However, more rigorous NFT studies are required in order to overcome the methodological weaknesses that we encountered in our qualitative analysis.

KEYWORDS

brain oscillatory activity, electroencephalogram, executive functions, neurofeedback training, task-switching, working memory

1 | INTRODUCTION

NeuroFeedback Training (NFT) is a re-emerging and promising brain training technique, consisting of a noninvasive neurophysiologically based method that allows individuals to learn to control and modulate their own brain activity (Angelakis et al., 2007; Doppelmayr & Weber, 2011; Enriquez-Geppert et al., 2013, 2017; Enriquez-Geppert, Huster, Scharfenort, et al., 2013; Jirayucharoensak et al., 2019; Ros et al., 2014). More specifically, during the training process, the neurofeedback system estimates the state of specific neural parameters and a computational interface provides trainees with continuous and real-time information about their physiological brain activity (e.g., employing video and/or audio signals), requiring them to self-regulate the neural parameter(s) and providing feedback that indicates whether the training goal is being achieved or not (Campos da Paz et al., 2018; Corydon Hammond et al., 2011; Enriquez-Geppert, Huster, & Herrmann, 2013; Jirayucharoensak et al., 2019). Thus, the trainee could successfully learn to modulate her/his brain activity through operant conditioning and/or modification of individual's self-perception (Lacroix, 1986), with possible beneficial effects on behavioral performance (Egner & Gruzelier, 2001; Engelbregt et al., 2016).

Moreover, NFT dynamically modulates brain activity as the individual can be trained to increase target frequencies (e.g., their rhythm or amplitude) or the activity of target brain areas and, at the same time, to inhibit other target frequencies/brain areas (Campos da Paz et al., 2018).

Different NFT approaches exist, and various parameters can be targeted during the training. Electroencephalographic (EEG) oscillations have been found to have a relationship with cognition and behavior. Groppe and colleagues (2013), for example, characterized the most common oscillations in the electrocorticogram, providing evidence for their function. Specifically, they suggested that alpha activity is related to sensory processing and attention, theta has a general role in cortical processing (e.g., top-down processing), and beta is involved in sensorimotor functions. Based on this association between different EEG frequencies and a variety of cognitive functions, the so-called *frequency-to-function mapping* (e.g., Fingelkurts & Fingelkurts, 2014), several NFT studies used selected features of electrical brain activity as the training parameter, with the aim of upregulating and/or downregulating specific endogenous neural oscillations related to precise cognitive functions (Enriquez-Geppert, Huster, Scharfenort, et al., 2014; Omejc et al., 2019). Furthermore, the rationale for EEG-NFT relies also on the evidence that different EEG parameters (e.g., frequency and/or amplitude) can be trained (Egner & Gruzelier, 2001; Hanslmayr et al., 2005; Zoefel et al., 2011). Consequently, EEG has become the most used NFT technique because it has low set-up cost (Escolano et al., 2011) and the modulation of neural oscillations using EEG-NFT has been shown to be effective for different frequencies associated with diverse cognitive processes (Enrique-Geppert et al., 2017; Gruzelier, 2014). Although it is also possible to implement NFT with other techniques such as fMRI by modulating the BOLD response (e.g., Zhang et al., 2013), the present systematic review will however cover only EEG-NFT approaches.

In the literature, many studies have used NFT as a therapeutic tool with clinical populations suffering from neurological and psychological disorders to normalize abnormal electrical oscillatory activity underlying various types of symptoms. Nevertheless, this review will not deal with clinical applications as they have already been extensively covered elsewhere. To cite some instances, there are reviews on psychiatric disorders in general (e.g., Arns et al., 2017; Micoulaud-Franchi et al., 2015), ADHD (e.g., Arns et al., 2009; Lofthouse et al., 2012; Yan et al., 2019), epilepsy (e.g., Nigro, 2019; Tan et al., 2009), and autism spectrum disorder (Coben et al., 2010).

Based on the NFT potential role in mediating cognitive and behavioral effects, this approach has also been applied to healthy individuals, outside of clinical research field, for boosting their behavioral performance and cognitive functioning, defined as "optimal" or "peak performance" (Egner & Gruzelier, 2001; Gruzelier, 2014; Vernon, 2005). This field of research is gaining increasing attention (Angelakis et al., 2007; Corydon Hammond et al., 2011; Enriquez-Geppert, Huster, & Herrmann, 2013; Jurewicz et al., 2018). EEG-NFT efficacy with healthy participants has however received criticism concerning the reliability of its effects as, to date, most of the works did not manage to provide evidence for unambiguous changes in behavioral and electrophysiological measures, especially due to methodological weaknesses, such as the lack of a sham/control group (Egner et al., 2004; Rogala et al., 2016; de Zambotti et al., 2012). In addition, subjects of controversy are also whether evidence from NFT studies with clinical populations can be applied to healthy individuals (Doppelmayr & Weber, 2011) and which indices should be adopted to quantify training success (Dempster & Vernon, 2009). Therefore, researchers point out the need of shared and rigorous methodological standards to overcome the scarcity of well-controlled studies and the heterogeneity of electrophysiological data (Egner et al., 2004; Ros et al., 2020).

A recent attempt to assess the state-of-the-art of EEG-NFT research on cognitive and affective outcomes in healthy individuals was made by Gruzelier (2014), who conducted a review of EEG-NFT, specifically concerning performance optimization and excluding clinical samples. The author assessed the effectiveness of different NFT protocols and found significant evidence for outcome gains and learning indices in several cognitive and affective domains. Rogala and colleagues (2016) tried to overcome Gruzelier's (2014) main limitation, namely, the inclusion of multiple studies with no proper control groups, by quantitatively assessing the efficacy of various EEG-NFT protocols to induce electrophysiological and behavioral changes, specifically focusing on attention and memory. Restricting the review only to well-controlled studies, they did not find evidence supporting a positive relationship between frequency band(s) changes and specific behavioral gains. Therefore, it is still debated whether and how NFT promotes healthy individuals' cognitive performance improvements (Doppelmayr & Weber, 2011). Even more controversial is whether NFT is an effective method for specifically enhancing performance on tasks tapping on executive functions, and, to the best of our knowledge, no systematic review has been previously focused on this particular topic (Enriquez-Geppert, Huster, & Herrmann, 2013; Rogala et al., 2016).

Executive functions are referred to as higher-order cognitive processes that enable, for instance, to flexibly set-up, regulate, and monitor goal-directed behaviors and thoughts by controlling lower-level cognitive operations, especially in novel or complex circumstances (MacPherson et al., 2019;

Miller & Cohen, 2001; Vallesi, 2020). Despite significant implications of executive functions for everyday life and their central role in human cognition and action regulation (Mischel et al., 2011; Miyake & Friedman, 2012), a limited extent of studies has targeted them using EEG-NFT (Enriquez-Geppert, Huster, & Herrmann, 2013). Furthermore, the few studies that implemented EEG-NFT to enhance healthy adults' executive functions had to deal with the issue of their multifaceted nature, which has led to a great variety of operational definitions of this construct (Barkley, 2012; Miyake & Friedman, 2012). Consequently, these studies adopted different definitions and used diverse tasks to measure them, producing a large methodological heterogeneity.

Therefore, the aim of the present systematic review was to assess whether, despite these limitations, existing EEG-NFT studies targeting executive functions provide reliable evidence for NFT effectiveness, both at the electrophysiological and at the behavioral level. We focused on EEG-NFT effects on healthy adults with no further age limitation, as we decided to include also studies involving healthy older adults to assess possible age-related differences.

Indeed, many studies provided evidence for an age-related decline in executive functions, among other domains, showing that healthy older adults perform poorer than the younger counterparts on executive function tasks documenting for instance working memory deficits, reduced inhibitory control, and decreased task-switching ability (MacPherson et al., 2015; West, 1996; Zanto & Gazzaley, 2019; but see Vallesi et al., 2021). As EEG-NFT could be a potentially effective method to counteract this age-related executive function decline, we assessed whether, to date, studies have provided reliable evidence for its effectiveness. To the best of our knowledge, there is no systematic review addressing specifically EEG-NFT and healthy older adults' executive functions, although there are some pieces of evidence that this technique could be successfully applied to older individuals (e.g., Angelakis et al., 2007; Gruzelier, 2014; Wang & Hsieh, 2013). As the number of retrieved articles dealing with healthy aging was limited, we could no perform age comparisons, but we discussed them in separate sections.

To formulate our research question, we adopted the PICOS approach (Liberati et al., 2009). Accordingly, the systematic review concerned controlled studies that performed EEG-NFT with healthy participants, including younger adults, older adults or both, with the aim of enhancing executive functions. Moreover, the present review dealt only with studies that assessed NFT efficacy both at behavioral and electrophysiological levels by comparing the experimental group(s) with a control group (see Method for detailed inclusion criteria).

Lastly, the present systematic review dealt only with a qualitative synthesis of the literature. A quantitative meta-analysis, although desirable, was in fact unfeasible, due to the limited number of retrieved studies for each given NFT protocol and specific executive function. Moreover, these studies used heterogeneous protocols and behavioral/electrophysiological efficacy indices to assess NFT effects, thus, not allowing quantitative comparisons.

In what follows, we shall briefly review evidence of the EEG frequency bands most often associated with executive functions before going to the method section. Specifically, we will discuss the rationale according to which theta, alpha, and low beta could represent potential NFT targets to enhance executive functions. We will focus on these specific bands as they are those used in the studies included in our systematic review (see Method for more details on our search and selection processes).

1.1 | Theta

Theta band power has been frequently related to performance on executive functions. In particular, a review by Klimesch (1999) provided evidence that a power increase in the theta band is positively associated with working memory load (also see Gevins et al., 1997; Grunwald et al., 2001; Jensen & Tesche, 2002).

This initial evidence was expanded by Cavanagh et al. (2012), who specifically investigated the involvement of theta recorded over fronto-medial brain regions (frontal-midline theta) during the execution of a variety of executive function tasks related to action monitoring, that is, when executive functions were required to integrate relevant information and to control action selection. Their findings supported the role of medial prefrontal cortex in reactive control, which in turn is reflected by frontal-midline theta activity, providing convincing evidence that this type of neural oscillation is involved in conflict monitoring and flexible behavior adjustments (Cavanagh et al., 2012).

In line with these findings, further studies provided evidence for the association of frontal-midline theta and executive functions. For example, the review by Mitchell et al. (2008) highlighted a clear relationship between working memory and frontal-midline theta activity. Furthermore, Nigbur et al. (2011) investigated whether frontal-midline theta was a marker for increased cognitive control during classical interference paradigms, such as the Simon, Flanker task, and Go/Nogo tasks. Overall, they found that theta power increased during the interfering conditions, indicating that theta reflected interference control (also see Cohen & Donner, 2013).

Overall, the above reviewed evidence suggests that the theta band represents a potential target-frequency for enhancing executive functioning.

1.2 | Alpha

Alpha rhythm is one of the dominant EEG phenomena in the human brain (Berger, 1930), and its activity has been correlated with several cognitive functions at all ages (Angelakis et al., 2007; Hanslmayr et al., 2005; Klimesch, 1999). Of interest to the current review, task-related alpha frequency has been associated with working memory and inhibitory control (Cooper et al., 2003; Klimesch et al., 1999, 2007). Accordingly, Klimesch and coworkers' (2007) review provided evidence that alpha event-related synchronization reflects inhibitory control, whereas the event-related desynchronization is involved in the gradual release of inhibition. Another review by Freunberger et al. (2011) proposed that alpha, by increasing signal-to-noise ratio, reduces interference from conflicting sensory stimuli, supporting working memory.

However, to minimize interindividual variability, Klimesch (1999) suggested to individually adjust the frequency window of alpha for each NFT trainee, by using Individual Alpha Frequency (IAF) as an anchor point. This allows, for instance, to target upper alpha band (i.e., the band 2 Hz above IAF), as it correlates with general cognitive performance (see also Escolano et al., 2011; Zoefel et al., 2011). Several NFT studies adopt this individualized approach (e.g., Angelakis et al., 2007; Hanslmayr et al., 2005; Klimesch et al., 1999). For instance, Mahjoory et al. (2019) suggested that resting state IAF reflects the activity of cognitive control networks at rest which subsequently sustain phasic performance.

1.3 | Low beta

Low beta activity in the 12- to 20-Hz frequency range enhances the signal-to-noise ratio in stimulus processing (Gruzelier, 2014) and sensorimotor rhythm (SMR) in the 12- to 15-Hz range is involved in the inhibition of sensory-motor cortex and, hence, response inhibition (Sterman, 1996). Moreover, low beta band and SMR were found to be associated with decreased impulsivity in clinical studies aiming at reducing these frequencies in individuals with hyperactivity and/or impulsivity disorders and showing symptom improvements (e.g., Thompson & Thompson, 2003). Accordingly, in a NFT study, Egner and Gruzelier (2001) showed that SMR enhancement over sensorimotor cortex reduced impulsiveness (i.e., fewer commission errors), improved response inhibition, and increased integration of relevant information.

Therefore, of interest to our review, SMR is assumed to be associated with thalamic inhibitory mechanisms by reducing sensorimotor interference and, consequently, improving cognitive performance (Kober et al., 2017; Sterman, 1996). Although there is evidence for low beta and SMR association with attention and memory (Gruzelier, 2014), the relation between these frequencies and executive functions is less clear. However, we retrieved several studies implementing low beta and SMR NFT including tasks tapping on executive functions.

2 | METHOD

2.1 | Eligibility criteria

Inclusion criteria were as follows:

1. *Population:* we included studies that used NFT with healthy adults and excluded those focused on clinical populations. Moreover, we reviewed also studies recruiting healthy older adults that had been prescreened to exclude pathological cognitive decline. We did not use a fixed age range to define the experimental samples and to classify them as younger or older adults, as we relied on the definition provided by the included studies. More details on age distributions will be given below, divided by NFT protocol. Overall, younger adults' age ranged from 20.7 to 46.4 years (*mean* = 24.69, *SD* = 6.37), whereas older participants' age ranged from 64.8 to 69.05 (*mean* = 67, *SD* = 1.95). NFT effects on younger and older adults were considered separately. More in detail, regarding studies that compared younger and older adults, we first analyzed NFT effects on younger participants, including those studies in the section dedicated to younger participants only. A separate subsection specifically concerned healthy older adults, including either studies that used only older participants or studies that compared older participants with younger ones.
2. *Intervention:* we reviewed studies that used EEG-NFT with the explicit purpose of enhancing executive functions, referring to them using the following terms: executive functions, cognitive control, or cognitive performance (in this case, the retrieved studies were included only if the authors used at least one executive function task in the battery). We additionally searched for terms concerning more specific executive functions that are commonly studied: task-switching, memory updating, response inhibition, conflict monitoring, working memory.

Since there is evidence of associations between specific EEG frequency band(s) and performance on particular cognitive tasks, we included studies that investigated whether individuals trained to enhance a particular EEG frequency, assumed to be involved in executive functions, and exhibited an improvement in the executive function task(s) (Vernon et al., 2003). In the literature, different EEG frequencies have been found to be involved in executive function tasks. Thus, various EEG-NFT protocols could be potentially effective in enhancing healthy individuals' executive functions. Given this heterogeneity, our initial database search was extended to all types of protocols that satisfied our research question. After screening all the retrieved records and assessing the eligibility of the articles, the studies that met our inclusion criteria were divided according to the EEG-NFT protocol used, namely, according to the target EEG frequency trained, resulting in a total of three NFT protocols (theta, alpha, and low beta). Each of these protocols included more than five studies and was discussed in a dedicated section. Moreover, when the retrieved article included more than one NFT protocol, we discussed each of them separately in the appropriate sections.

3. *Comparator*: we included only studies that compared an experimental group undergoing NFT to modulate a target-frequency with a control group. The control group might have been passive, sham, or a group that received a different NFT protocol (e.g., training of EEG-frequency different from the target one). More in detail, as passive groups, we considered control groups that underwent pretraining and posttraining assessment in the same interval of time as the experimental group but did not participate in intermediate sessions. Of note, this type of control has many limitations and only allows to control for the impact of practice effects of the behavioral assessment, without taking into account possible placebo and social effects arising from the experimental setting (Zoefel et al., 2011). Thus, if the study comprised more than one experimental group, we focused on one of the experimental groups as the control group. The following alternatives were considered as control groups with a sham NFT: (i) groups that had the same design as the NFT group but received NFT only for the first session whereas in the subsequent sessions received a replayed feedback of the first session, (ii) groups that received random feedback or feedback from a frequency that was not the target-frequency of the study, and (iii) pseudo-NFT groups that received a playback feedback of other NFT group participants (but real eyeblink activity). Another alternative, which however could not be considered as sham, were groups that did not receive NFT at all but were engaged in the same number of sessions with the same duration of the NFT group with a control activity (e.g., while undergoing a behavioral training), thus having the same amount of experimenter-contact as the experimental group. A control group just differing in the trained frequency band was of course one of the most desirable options, enabling to control many interfering variables (Zoefel et al., 2011). Lastly, if the study did not include a control group but comprised more than one experimental group, we evaluated whether it was possible to use one of the experimental groups as control, assuming that the groups could serve as controls for one another (e.g., Rogala et al., 2016).
4. *Outcomes:* NFT efficacy was assessed at both the electrophysiological and behavioral levels since significant ($p < .05$) changes in the EEG activity and in the target cognitive functions may provide evidence for the validity of EEG-NFT protocols (Rogala et al., 2016). Specifically, we verified whether the experimental group(s) significantly differed from the control one(s) on each of the two posttraining types of measures (see Definition of successful training for further details). Therefore, to gain insight about the overall efficacy of the training, we included only studies that provided both electrophysiological and behavioral measurements before and after NFT.

As electrophysiological measures we considered resting-state EEG before and after NFT and/or real-time EEG changes within and across sessions when available. In addition, we included both studies that assessed NFT effects on continuous EEG, namely, as resting or passive baseline measurements and studies that calculated NFT effects on EEG rhythms assessed during active tasks.

Concerning behavioral measures, we included studies that assessed performance on executive function tasks, before and after the NFT training. We did not decide a priori which executive function tasks could be incorporated in the systematic review, as we used a permissive inclusion criterion, that is, included studies had to encompass at least one (or more) task tapping executive functions. When the studies used neuropsychological batteries, we evaluated whether at least one of the neuropsychological tests or indices measured executive functions.

2.2 | Information sources and search

Articles were identified by searching three electronic databases (i.e., Pubmed, PsychInfo, Scopus) and by examining the retrieved items. The keywords were the following: EEG neurofeedback AND healthy adults/healthy young adults/normal subjects AND executive functions/cognitive control/executive control/executive processes AND enhancement/improvement/boosting.

Through database searching, 711 records were identified, whereas 2 records were identified through other sources (i.e., examination of the reference lists of the retrieved articles).

After removing duplicates, 304 unique articles remained (see Figure 1).

2.3 | Selection process

First, we screened all the retrieved records by evaluating their titles and abstracts. We selected works using NFT, recruiting healthy younger and/or older adults and referring

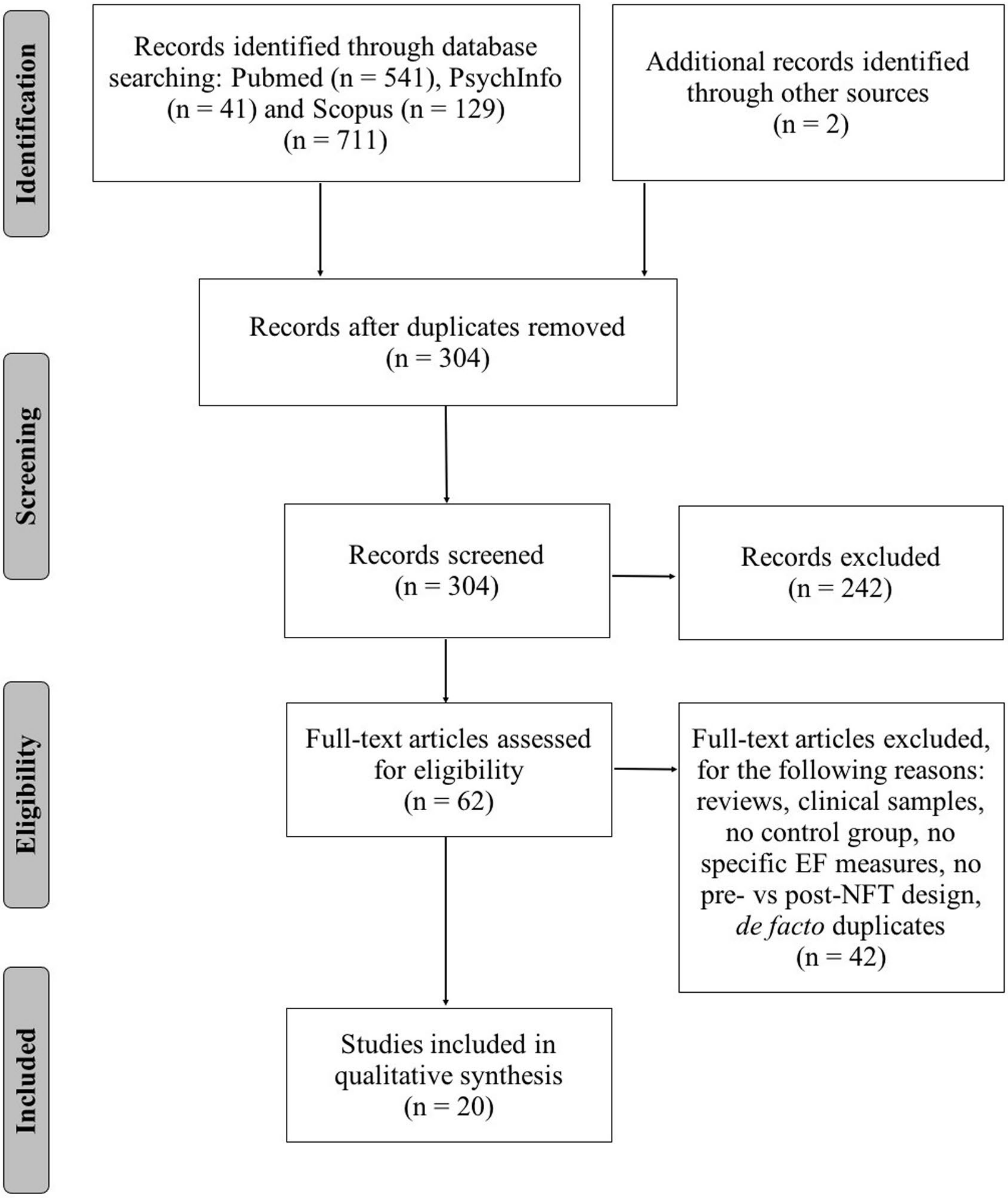


FIGURE 1 PRISMA flow diagram of the studies screened, assessed for eligibility and included in the review

to executive function/cognitive control/executive control/executive processes. In this initial screening phase, we excluded articles whose titles and abstracts indicated that our inclusion criteria were not met. More in detail, we excluded 111 articles because they involved rehabilitation programs (e.g., motor rehabilitation) and targeted clinical populations (e.g., ADHD, autism, and depression). Moreover, 37 articles were excluded as they did not use NFT but other stimulation techniques (e.g., TMS, tDCS), whereas 23 were excluded because NFT was used with techniques other than EEG (e.g., fMRI). Other articles did not meet our criteria as NFT was not used to specifically train executive functions but other abilities ($n = 55$; e.g., motor imagery, motor performance, and emotional self-regulation) or since behavioral measures were not collected ($n = 14$). Lastly, 1 article dealt with aromatherapy and 1 article with healthy children. Overall, with this screening 242 articles were excluded.

Thus, we identified 62 eligible publications, whose full-text was then assessed. This second selection process included papers that met our inclusion criteria. We excluded publications which did not involve experimental studies (e.g., reviews, $n = 8$), included also clinical samples and did not allow us to use only the data collected on the healthy sample ($n = 5$), did not include a control group (and we could not consider a second experimental group as control, $n = 5$), did not measure specifically executive functions (e.g., mental rotation tasks, attention tasks and so on, $n = 19$), and did not have a preversus post-NFT design ($n = 2$) and de facto duplicates among which a corrigendum of another article ($n = 1$) and conference papers ($n = 2$).

The final set of articles selected for the qualitative synthesis was composed of 20 publications, as shown in the PRISMA flow diagram in Figure 1.

2.4 | Grouping and description of included studies

During the selection process, we found marked differences between studies, especially at the methodological level. Therefore, we applied restrictive criteria in order to make our dataset as homogenous as possible. In addition to the unresolved issue regarding the efficacy of EEG-NFT on executive functions, the identification of an oscillatory activity candidate involved in this specific cognitive process is of equal importance (Enriquez-Geppert, Huster, & Herrmann, 2013). In this regard, since there is no consensus on which frequency is better to use as target, a wide variety of NFT protocols has been used by the selected papers. To address the issue of the large heterogeneity of the included studies, we divided them into groups based on (i) NFT protocol used (EEG-frequency trained and NFT direction) and (ii) target-EF (Table 1).

This review is organized in three main sections according to the target-frequency: theta (4–8 Hz), alpha (8–12 Hz), and low beta (including only studies targeting 12–20 Hz). Both

TABLE 1 EEG-NFT studies divided according to training protocols

		Target executive-function			
Protocols	Study	Task-switching	Response inhibition	Conflict monitoring	Working memory and memory updating
Theta (4–8 Hz)	Enriquez-Geppert, Huster, Figge, et al. (2014)	Yes	Yes	Yes	Yes
	Gonçavales et al. (2018)	No	No	Yes	No
	Reis et al. (2016)[a]	No	Yes	Yes	Yes
	Vasquez et al. (2015)	No	Yes	No	No
	Vernon et al. (2003)	No	No	No	Yes
	Wang and Hsieh (2013)[b]	No	No	Yes	No
	Xiong et al. (2014)	No	No	No	Yes
Alpha (8–12 Hz)	Berger and Davelaar (2018)	No	Yes	Yes	No
	Escolano et al. (2011)	No	No	No	Yes
	Escolano et al. (2014)	No	Yes	Yes	Yes
	Gomez-Pilar et al. (2016)[a]	No	Yes	No	No
	Gordon et al. (2020)	Yes	Yes	No	Yes
	Hsueh et al. (2016)	No	No	No	Yes
	Naas et al. (2019)	No	No	No	Yes
	Pei et al. (2018)	No	No	Yes	Yes
	Reis et al. (2016)[a]	No	Yes	Yes	Yes
	Wei et al. (2017)	No	No	No	Yes
Beta (only 12–20 Hz)	Campos da Paz et al. (2018)[a]	No	No	No	Yes
	Cannon et al. (2009)	No	No	No	Yes
	Egner and Gruzelier (2004)	No	Yes	No	No
	Gomez-Pilar et al. (2016)[a]	No	Yes	No	No
	Gonçalves et al. (2018)	No	No	Yes	No
	Kober et al. (2017)	No	No	No	Yes
	Vasquez et al. (2015)	No	Yes	No	No
	Vernon et al. (2003)	No	No	No	Yes

Note: Target EF lists the EFs targeted in each study.

Abbreviations: EEG-NFT, electroencephalographic-neurofeedback training.

[a] Studies including older adults only.

[b] Studies including both younger and older adults. Studies that included only younger adults were not marked with asterisks.

upregulation and downregulation of the band were included in the same protocol section (with the direction of band regulation appropriately specified and discussed). Therefore, we also addressed the issue of whether one of these three specific frequencies plays a more critical role than the others in the modulation of executive functions.

Specific sections based on target executive functions were not created due to the low number of papers for each of these functions. Of note, many studies targeted more than one executive function at a time and could not be included exclusively in one set. Overall, the selected studies used at least one (or more) of the following tasks: anti-saccade task, Attention Network Test – ANT (conflict index), backward digit span, conceptual span, Corsi Block Tapping Test (CBTT) backwards task, delayed matched-to-sample task, Go/Nogo, Iowa Gambling task, Luria-AND test (attentional control test), n-back, operation span, Paced Auditory Serial Addition Task (PASAT), modified Sternberg recognition task, stop-signal, Stroop, task-switching, Trail Making, Working Memory Index (WMI) of the WAIS-III. Lastly, for each protocol section, one sub-section concerned specifically executive function enhancement in older participants.

2.5 | Definition of successful training

According to Gruzelier (2014), there are two sources of evidence that can be used to assess NFT efficacy: (i) when the NFT experimental group shows a successful cognitive outcome, namely better cognitive performance on behavioral tasks, compared with a control group and/or comparison condition and (ii) when there is evidence of NFT learning in the posttraining assessment. Therefore, NFT was considered as successful in enhancing executive functions when behavioral and electrophysiological effects were both clearly shown. More in detail, even if the comparison between pretraining and posttraining measures, performed for each group separately (e.g., paired-samples *t*-tests), reached significance in the experimental group but not in the control group, we did not consider this as sufficient evidence for the efficacy of the training unless a significant Training phase by Group interaction was also reported (see Nieuwenhuis et al., 2011), both for behavioral and electrophysiological outcomes. Additionally, to better clarify the significant interactions, we verified whether the training effect on the NFT group was in the intended direction (e.g., greater amplitude of a frequency band if the direction of NFT was upregulation and performance enhancement in NFT group). Lastly, we reported the statistics used by each study and the effect sizes. When multiple comparisons were performed, we checked whether the measures had been corrected. If not, we reported and discussed this issue in the result section.

On the basis of these criteria, we used a binary approach, that is, experiments were qualified as "successful" when NFT produced both significant behavioral and electrophysiological effects, whereas they were considered as "not successful" when NFT did not produce significant behavioral and/or electrophysiological effects. Thus, in line with Rogala and colleagues (2016), we calculated the success ratio (SR), defined as the percentage of successful studies out of the total number of studies included in each protocol type.

In the event that NFT outcomes were only partially successful, that is, either at the behavioral level or at the electrophysiological one, we discussed possible study weaknesses underlying the partial success and classified them as "partially successful". Of note, these studies were not considered as "successful" for the calculation of the SR.

2.6 | Theta protocols

Seven of the retrieved articles used theta protocols: five included younger adults only, one older adults only, and one both younger and older adults. As shown in Table 2, the most commonly used electrode locations were frontal and central. Of note, the distinction between frontal-midline theta and theta in general regarded mainly the terminology. Concerning older adults theta protocols, the first of the two retrieved studies trained frontal-midline theta measured from Fz, whereas the second targeted more general theta measured from Fp1, Fp2, Fz, Pz (Table 3).

2.7 | Alpha protocols

Ten of the of the retrieved articles used alpha protocols: eight with younger adults only and two with older adults only. Tables 4 and 5 provide a detailed description of the characteristics of the NFT studies with younger and older adults, respectively.

2.8 | Low beta protocols

Eight of the retrieved articles used low beta protocols, among which one article included 2 experiments. Therefore, we analyzed nine studies: seven with younger adults only and two with older adults only. Among low beta protocols with younger adults (Table 6), two studies trained beta in the 15- to 18-Hz frequency range and measured beta from Cz ($n = 1$) or from 19 leads ($n = 1$), whereas one experiment targeted beta in the 13- to 21-Hz range over Cz ($n = 1$). In addition, four protocols trained specifically the SMR (12–15 Hz) and used electrodes located over Cz ($n = 4$).

Out of two low beta studies with older adults, one trained beta band in the range 18–21 Hz over C3, Cz, C4, while the other targeted SMR measured from Cz (Table 7).

TABLE 2 List of studies using theta neurofeedback protocols with younger adults and main study characteristics

Study	Target frequency	Electrodes position	Target EF(s)	Direction of NFT	Single versus multiband	Number of NFT sessions	Total minutes of NFT	Sample size	Average age	Control group type
Enriquez-Geppert, Huster, Figge, et al. (2014)	fm-theta	Fz, FC1, FC2 FCz, Cz	Task-switching, response inhibition, conflict monitoring, WM	Up	Single	8	240	40	24.8	Sham pseudo-NFT (receiving playback feedback from NFT group)
Gonçalves et al. (2018)	theta	Cz	Conflict monitoring	Up	Multi (SMR-)	1	25	30	20.7	Opposite experimental protocol (SMR+, theta -)
Vasquez et al. (2015)	theta	Cz	Response inhibition	Down	Multi (beta+)	1	30	30	23.4	Passive
Vernon et al. (2003)	theta	Cz	WM	Up	Multi (delta- and alpha-)	8	120	30	22.1	Different experimental protocol (SMR+, theta-, beta-)
Xiong et al. (2014)	theta	Fz, FCz, Cz, C1, C2	WM	Up	Multi (alpha-)	5	10	48	Not reported	3 control groups: sham random NFT, non-training, behavior-training
Wang and Hsieh (2013)	fm-theta	Fz	Conflict monitoring	Up	Single	12	180	16	22.2	Sham NFT (to enhance a randomly selected frequency)

Abbreviation: NFT, neurofeedback training.

TABLE 3 List of studies using theta neurofeedback protocols with older adults and main study characteristics

Study	Target frequency	Electrodes position	Target EF(s)	Direction of NFT	Single versus multiband	Number of NFT sessions	Total minutes of NFT	Sample size	Average age	Control group type
Reis et al. (2016)	theta	Fp1, Fp2, Fz, Pz	WM	Up	Multi (alpha+)	4	120	34	65.97	3 control groups: Sham NFT, NFT+ cognitive training, only cognitive training
Wang and Hsieh (2013)	fm-theta	Fz	Conflict monitoring	Up	Single	12	180	16	64.8	Sham NFT (to enhance a randomly selected frequency)

Abbreviation: NFT, neurofeedback training.

TABLE 4 List of studies using alpha neurofeedback protocols with younger adults and main study characteristics

Study	Target frequency	Electrodes position	Target EF(s)	Direction of NFT	Single versus multiband	Number of NFT sessions	Total minutes of NFT	Sample size	Average age	Control group type
Berger and Davelaar (2018)	Alpha	Fp2	Response inhibition, conflict monitoring	Up	Single	5	125	22	35.2	Different experimental group (2D NFT group)
Escolano et al. (2011)	Upper alpha	P3, Pz, P4, O1, O2	WM	Up	Single	5	125	16	24.7	Passive (behavioral measures only)
Escolano et al. (2014)	Individual upper alpha	P3, Pz, P4, O1, O2	Response inhibition, conflict monitoring, WM	Up	Single	1	25	19	25.05	Sham NFT (receiving playback feedback from NFT group)
Gordon et al. (2020)	Individual upper alpha	Pz	Task-switching, response inhibition, WM	Up	Single	10	150	165	22.12	3 control groups: WM training, active control training, silent control group
Hsueh et al. (2016)	Alpha	C3, Cz, C4	WM	Up	Single	12	432	50	21.3	Sham NFT (feedback of randomly selected frequency)
Naas et al. (2019)	Individual upper alpha	P7, O1, O2, P8	WM	Up	Single	4	60	33	21.27	Sham NFT
Pei et al. (2018)	Alpha	Fz, C4	Conflict monitoring, WM	Up	Single	5	180	20		Sham NFT (feedback of randomly selected frequency)
Wei et al. (2017)	Alpha	C3	WM	Up	Single	12	300	30	26	Sham NFT (feedback of randomly selected frequency)

Abbreviation: NFT, neurofeedback training.

TABLE 5 List of studies using alpha neurofeedback protocols with older adults and main study characteristics

Study	Target frequency	Electrodes position	Target EF(s)	Direction of NFT	Single versus multiband	Number of NFT sessions	Total minutes of NFT	Sample size	Average age	Control group type
Reis et al. (2016)	alpha	Fp1, Fp2, Fz, Pz	Response inhibition, conflict monitoring, WM	Up	Multi (theta+)	4	120	34	65.97	3 control groups: Sham NFT, NFT+ cognitive training, only cognitive training
Gomez-Pilar et al. (2016)	alpha	C3, Cz, C4	Attentional control, response inhibition	Down	Multi (beta+)	5	450	63	68.15	Passive (behavioral measures only)

Abbreviation: NFT, neurofeedback training.

TABLE 6 List of studies using low beta neurofeedback protocols with younger adults and main study characteristics

Study	Target frequency	Electrodes position	Target EF(s)	Direction of NFT	Single versus multiband	Number of NFT sessions	Total minutes of NFT	Sample size	Average age	Control group type
Cannon et al. (2009)	Low beta (14–18)	LNFB conducted using 19 leads (FP1, FP2, F3, F4, Fz, F7, F8, C3, C4, Cz, T3, T4, T5, P3, P4, Pz, O1, O2	WM	Up	Single	33	528	14	21.21	Different experimental groups that received NFT from lDLPFC and rDLPFC
Egner and Gruzelier (2004)	Beta 1 (15–18)	Cz	Response inhibition	Up	Single	10	150	16	21.7	Behavioral training
Egner and Gruzelier (2004)	SMR (12–15)	Cz	Response inhibition	Up	Single	10	150	17	21.7	Behavioral training
Gonçalves et al. (2018)	SMR	Cz	Conflict monitoring	Up	Multi (theta-)	1	25	30	20.7	Opposite experimental protocol (theta+, SMR-)
Kober et al. (2017)	SMR	Cz	WM	Up	Single	10	450	20	46.4	Different experimental protocol (gamma+)
Vasquez et al. (2015)	beta 13–21	Cz	Response inhibition	Up	Multi (theta-)	1	30	30	23.4	Passive
Vernon et al. (2003)	SMR	Cz	WM	Up	Multi (theta- and beta-)	8	120	30	22.1	Different experimental protocol (theta+, delta-, alpha-)

TABLE 7 List of studies using low beta neurofeedback protocols with older adults and main study characteristics

Study	Target frequency	Electrodes position	Target EF(s)	Direction of NFT	Single versus multiband	Number of NFT sessions	Total minutes of NFT	Sample size	Average age	Control group type
Campos da Paz et al. (2018)	SMR	Cz	WM	Up	Single	10	90	17	69.05	2 control groups: sham NFT (real NF only first session), No-NFT control group
Gomez-Pilar et al. (2016)	beta (18–21)	C3, Cz, C4	attentional control, response inhibition	Up	Multi (alpha-)	5	450	63	68.15	Passive (behavioral measures only)

Abbreviation: NFT, neurofeedback training.

3 | RESULTS

In the following sections, we will summarize and discuss the review results about the different protocols. The sections are organized according to the targeted band (theta, alpha, low beta) and the population being trained (younger adults, older adults). Each section will be accompanied by a summary table in which statistics and p-values for each target-measure will be reported. Specifically, we will outline the Training phase by Group interaction, upon which we based our judgment of NFT success. Furthermore, we will highlight the specific effect of training on the experimental group. Lastly, the tables will contain the effect sizes, which were reported only in few studies and, thus, in most cases, were estimated by us.

3.1 | Theta protocols with younger adults

Theta NFT with younger adults obtained a SR of 33.33% (Table 8). The studies by Enriquez-Geppert et al. (2014) and Wang and Hsieh (2013) were qualified as "successful", as they yielded positive results in both electrophysiological and behavioral domains. Notably, these studies were the only ones among those using theta NFT protocols that targeted specifically frontal-midline theta and that used a single-band protocol; that is, both of them trained participants to upregulate frontal-midline theta irrespective of any other frequency. Enriquez-Geppert, Huster, Figge and collaborators (2014) used Fz, FC1, FC2, FCz, and Cz as the training electrodes and calculated the average frontal-midline theta activity, whereas Wang and Hsieh (2013) employed the Fz electrode only. Thus, we might speculate that the Fz electrode should be included to achieve NFT success in boosting executive functions. Moreover, these two experiments conducted a higher amount of NFT minutes compared to the "not successful" ones, 240 and 180 min, respectively. Regarding the control condition, both studies used sham NFT groups. However, while Enriquez-Geppert, Huster, Figge and coworkers (2014) employed a sham pseudo-NFT group in which control group participants received the feedback from another participant belonging to the experimental group, Wang and Hsieh (2013) provided control group participants with feedback of randomly selected frequency bands (10–13, 13–16, 16–20, or 20–25 Hz). Of note, the latter study examined age-related differences and, for this reason, it included two distinct NFT groups, one comprising younger and another one comprising older adults and two control groups, respectively. Regarding executive function assessment, these two studies used different behavioral tasks and obtained diverse results. In fact, Enriquez-Geppert, Huster, Figge and colleagues (2014) investigated NFT effects on behavioral performance using four executive function tasks. In the visual three-back task (measuring working memory), participants were presented with

TABLE 8 Statistics and results of theta neurofeedback protocols with younger adults

		EEG			Behavioral		
	Measure	Training × group interaction	Training effect on NFT group	Measure	Training × group interaction	Training effect on NFT group	NFT success
Enriquez-Geppert, Huster, Figge, et al. (2014)	ind_$_{\text{fm}}$Theta amplitude	**$_{\text{mixed}}F = 6.23$, $p < .001$, $d = 0.80$**	NA (NT)	3-back, ACC	**$_{\text{1t\_unpaired}}t = 1.79$, $p = .041$, $d = 0.57$**	**$_{\text{2t\_paired}}t = 4.71$, $p < .001$, $d = 1.08$**	**Successful**
				TS_Switch, RT	**$_{\text{1t\_unpaired}}t = 1.97$, $p = .028$, $d = 0.62$**	**$_{\text{2t\_paired}}t = 7.43$, $p < .001$, $d = 1.70$**	
				TS_Stay, RT	**$_{\text{1t\_unpaired}}t = 1.70$, $p = .049$, $d = 0.54$**	**$_{\text{2t\_paired}}t = 4.85$, $p < .001$, $d = 1.11$**	
				TS_Switch, ACC	$_{\text{1t\_unpaired}}t = 1.63$, $p = .055$, $d = 0.51$	**$_{\text{2t\_paired}}t = 2.70$, $p = .015$, d = 0.62**	
				TS_Stay, ACC	$_{\text{1t\_unpaired}}t = -1.65$, $p = .947$, $d = -0.52$	**$_{\text{2t\_paired}}t = 3.19$, $p = .005$, $d = 0.73$**	
				STR_Inc, RT	$_{\text{1t\_unpaired}}t = -0.25$, $p = .597$, $d = -0.08$	*$_{\text{2t\_paired}}t = 1.50$, $p = .151$, $d = 0.34$*	
				STR_Inc, ACC	$_{\text{1t\_unpaired}}t = 0.99$, $p = .164$, $d = 0.31$	**$_{\text{2t\_paired}}t = 3.76$, $p = .001$, $d = 0.86$**	
				SS_Stop, RT	$_{\text{1t\_unpaired}}t = -1.56$, $p = .936$, $d = -0.49$	$_{\text{2t\_paired}}t = 0.70$, $p = .489$, d = 0.16	
				SS_Go, RT	$_{\text{1t\_unpaired}}t = 1.65$, $p = .054$, $d = 0.52$	NA (NT)	
Gonçalves et al. (2018)	$_{\text{fm}}$Theta amplitude	$_{\text{mixed}}F = 3.47$, $p = .073$, $d = 0.68$	**$_{\text{2t\_paired}}t = 3.19$, $p = .001$, $d = 0.85$**	ANT_Conflict, RT	$_{\text{mixed}}F = 1.35$, $p = .255$, $d = 0.30$	$_{\text{2t\_paired}}t = 0.36$, $p = .728$, $d = 0.09$	Not Successful
Vasquez et al. (2015)	$_{\text{fl}}$Theta amplitude (-)	NA (NR)	$_{\text{2t\_paired}}t = 0.66$, $p = .515$, $d = 0.21$	IOWA	**$_{\text{mixed}}F = 3.68$, $p = .039$, d = 0.70**	**$_{\text{2t\_paired}}t = 2.52$, $p = .033$, $d = 0.80$**	Partially Successful
Vernon et al. (2003)	$_{\text{fm}}$Theta/Delta ratio	NA (NT)	$_{\text{2t\_paired}}t = 0.45$, $p = .663$, $d = 0.14$	Conceptual Span, ACC	**$_{\text{mixed}}F = 5.20$, $p = .004$, $d = 0.83$**	NA (NS, NR)	Not Successful
	$_{\text{fm}}$Theta/Alpha ratio	NA (NT)	$_{\text{2t\_paired}}t = 2.26$, $p = .054$, $d = 0.71$				
Wang and Hsieh (2013)	$_{\text{fm}}$Theta amplitude	**$_{\text{mixed}}F = 10.39$, $p = .003$, $d = 1.14$**	**$_{\text{2t\_paired}}t = 2.69$, $p = .031$, $d = 0.95$**	ANT_Conflict, RT	**$_{\text{mixed}}F = 32.59$, $p < .001$, $d = 2.02$**	**$_{\text{2t\_paired}}t = 8.17$, $p < .001$, $d = 2.58$**	**Successful**
				ANT_Conflict, ACC	NS (NR)	NA (NT)	
				WM	**$_{\text{mixed}}F = \text{NR}$, $p < .05$, $d = \text{NR}$**	NA (NS, NR)	
Xiong et al. (2014)	$_{\text{fm}}$Theta/Alpha ratio	NA (NT)	NA (NT)	2-Back, RT	NA (NT)	**$_{\text{2t\_paired}}t = 8.61$, $p = .001$, $d = 2.38$**	Not Successful
				2-Back, ACC	NA (NT)	**$_{\text{2t\_paired}}t = 5.32$, $p = .006$, $d = 3.85$**	

Note: Bold text indicates significant results; italic text indicates results that were not reported in the original article but could be estimated from available data; underlined text indicates reported results that we transformed for the sake of homogeneity.

Abbreviations: ANT, attention network task; $_{\text{fl}}$Theta, fronto-lateral theta; $_{\text{fm}}$Theta, frontal-midline theta; ind_$_{\text{fm}}$Theta, individual frontal-midline theta; NA, not available; NFT, neurofeedback training; NR, not reported; NS, not significant; NT, not tested; SS, stop-signal task; STR, Stroop task; TS, task-switching paradigm.

letter sequences and were asked to report whenever a letter had already been presented three trials before; in the number-letter task-switching, they viewed number-letter pairs and were instructed to classify either the numbers or the letters based on the specific background color; in the Stroop task, participants responded to the ink color of color-words (this task measured reactive control of interference for incongruent ink-word combinations); lastly, in the visual stop-signal task (measuring reactive response inhibition), they were instructed to abort their initiated responses to the direction arrows when those changed their color. The authors reported improvements on the three-back and task-switching tasks in both groups, but these changes were significant in the NFT group only. Indeed, independent-samples *t*-tests comparing the NFT group with the pseudo-NFT group on pretraining versus posttraining differences (which corresponds to a Training phase by Group interaction) reached significance for accuracy in the three-back condition of the three-back task and for RT in the switch and stay conditions on the letter-number task-switching. By contrast, no significant interaction effect was observed for the Stroop and the stop-signal tasks. Our estimates of training effects on the NFT group confirmed that the training induced changes in the intended direction, as indexed by the significant accuracy increase on the three-back task and the significant RT decrements in the switch and stay conditions on the task-switching. Hence, in this study, NFT was able to enhance proactive control indices (i.e., memory updating and mental set-shifting) but did not produce the desired effects on proxies of reactive control (i.e., conflict monitoring and motor inhibition). By contrast, Wang and Hsieh (2013) assessed conflict monitoring and working memory changes, using the Attention Network Test (ANT) and the modified Sternberg Recognition task, respectively. The former consisted in the presentation of a visual cue, followed by a central arrow flanked by four arrows, and the participants were required to indicate whether the central target arrow pointed in the same direction as the other four (congruent) or in the opposite direction (incongruent). From this task, conflict scores were calculated by subtracting the mean RT of all congruent conditions from the mean RT of the incongruent conditions. In the modified Sternberg Recognition task, instead, participants were presented with word lists and were asked to judge a subsequently shown probe word as “old” or “new”, according to whether they had been presented before or not. After NFT, participants demonstrated a significant performance enhancement with respect to pre-NFT measures and to the control group, as indexed by a significant interaction of Training phase and Group, driven by the reduction in conflict scores in the two NFT groups. Of note, this result refers to the aggregate of both younger and older adult participants and does not allow to disentangle age-related differential effects. Concerning the electrophysiological results, in Enriquez-Geppert, Huster, Figge and colleagues' (2014) study, NFT effects on frontal-midline theta were quantified as relative changes of amplitude, namely compared to the first training session values. The authors demonstrated a significant increase in frontal-midline theta amplitude, across sessions, and compared to the control group. Similarly, Wang and Hsieh (2013) found a significant interaction between Training and Group, which indicated that NFT effects on theta were specific and limited to the groups receiving NFT. This result was further strengthened by the evidence of a significant frontal-midline theta increase in younger adult NFT group. Overall, both Enriquez-Geppert, Huster, Figge and colleagues' (2014) and Wang and Hsieh's (2013) findings were characterized by medium to very large effect sizes.

Among theta protocols, three studies did not obtain significant results, either at the behavioral level or at the electrophysiological one. Gonçalves et al. (2018) used two multiband protocols during 25 min of a single session NFT, the former requiring participants to upregulate theta and downregulate SMR and the second involving the opposite training (SMR upregulation and theta downregulation). A paired-samples *t*-test, contrasting theta amplitude at baseline and during the last block, revealed a significant increase in theta after theta NFT. However, although the changes were in the intended direction, they were not supported by a significant Training phase by Group interaction. For this reason, we could not consider this protocol as effective at inducing electrophysiological effects. Similarly, this short NFT protocol did not enhance conflict monitoring, as shown by the unchanged conflict scores on the ANT after theta NFT and further confirmed by our Training phase by Group interaction estimate, which revealed no significance.

In Vernon and colleagues' (2003) study, theta group was required to upregulate theta while inhibiting delta and alpha and was compared to a SMR group, trained to upregulate SMR while downregulating theta and beta. In both cases the feedback was provided from the Cz electrode. Executive function performance was measured on a conceptual span task during which, after the presentation of words belonging to three different semantic categories, a cue indicated which category words participants had to recall. Albeit each participant underwent 120 min of NFT, theta group did not increase theta ratio from period 1 to period 5 and did not achieve any working memory enhancement. More in detail, concerning behavioral performance, a significant interaction between Training phase and Group was reported, but working memory increase was not significant for the theta NFT group, while at the electrophysiological level, the training did not affect either theta/delta or theta/alpha ratios.

Lastly, Xiong and colleagues (2014) implemented a multiband protocol to upregulate theta while suppressing alpha to assess effects on working memory, measured during a spatial two-back task, in which participants were required to judge whether the current stimulus was identical to the one shown two positions back in the sequence. The group receiving NFT was compared to a sham-NFT group, to a non-training group undergoing only pretest and posttests and to a behavioral-training group. They reported a better performance on the working memory task after 10 min of NFT, but their behavioral analyses suffer from the issue that only pairwise comparisons between pre-NFT and post-NFT for each group separately were performed, without directly testing any Training phase by Group interaction. Consequently, it remains to be demonstrated whether NFT group's working memory performance improved more than the control groups' performance. A further drawback concerns electrophysiological effects. Although the authors asserted that their protocol was effective at increasing theta/alpha ratio, they did not report any data supporting this conclusion. Hence, electrophysiological and behavioral effects were both regarded as not significant here.

The study by Vasquez et al. (2015) was classified as "partially successful". The 30-min single session of NFT, during which participants were required to downregulate theta while upregulating standard beta (13–21 Hz), did not induce changes in theta band but yielded positive results in the behavioral domain. More in detail, the authors reported a significant increase in response inhibition on the Iowa Gambling Task after NFT which, however, occurred only in the experimental group with the active electrode over Cz and the reference electrode in the right ear lobe (right hemisphere group). In fact, the Training phase by Group interaction was due to significant statistical changes in the posttraining phase in the right hemisphere group only.

3.2 | Theta protocols with older adults

Theta NFT with older adults succeeded for one out of two cases, which provided a SR of 50% (see Table 9). The "successful" study was the one by Wang and Hsieh (2013), which applied to older participants the same methodology described in the previous section with younger adults. Of note, frontal-midline theta before NFT showed age-related differences, with a decrease in theta amplitude in the fronto-central midline region of the scalp. However, despite this age-related decline, there was a significant Training phase by Group interaction, suggesting that NFT effects on theta were specific and limited to the groups receiving NFT. Moreover, the difference in frontal-midline theta amplitude between pretraining and posttraining was significant and in the intended direction in older participants receiving NFT. EEG changes induced also behavioral enhancements on the ANT conflict score and on the modified Sternberg Recognition task. After NFT, they found greater conflict monitoring in terms of lower conflict costs in both NFT groups. Nevertheless, these results do not distinguish specific behavioral effects on older adult NFT group. Additionally, after NFT, older adults, but not younger ones, improved their working memory performance as shown by the increased accuracy on the modified Sternberg recognition task.

By contrast, Reis and coworkers' (2016) study was "not successful". They implemented a multiband protocol which required participants to upregulate theta in the first four NFT sessions and to upregulate alpha in the last four NFT sessions, using Fp1, Fp2, Fz, and Pz as training electrodes. In addition to the NFT group, there were three control groups: a sham-NFT group, a group undergoing NFT for some blocks and then cognitive task blocks, in which the intensity of NFT was lower compared to the experimental group, and, lastly, a cognitive training group. EEG changes were tested only by comparing theta power spectrum density (PSD) before and after training in the NFT group, revealing a training effect on theta power during baseline and activity. However, the lack of a statistical comparison between groups prevents from confirming the efficacy of the NFT. In the posttraining behavioral assessment, no significant improvement was observed in the NFT group performance on the Backward Digit Span test, during which volunteers were asked to repeat the auditory presented digits but in the reverse order. Thus, this study did not provide evidence for an enhancement of older adults' working memory. In addition to the outlined statistical limitations, this study suffers from a methodological pitfall, since the combination of theta and alpha NFT on the same experimental group prevents from isolating the specific training effects of the two frequency bands.

3.3 | Alpha protocols with younger adults

Alpha NFT protocols with younger adults were quantitatively superior but resulted in a lower SR (12.5%). Out of eight studies, only one was "successful", four were "not successful", and three "partially successful" (Table 10).

The successful study was the one conducted by Berger and Davelaar (2018), who trained participants to increase alpha amplitude for 125 min of NFT placing the feedback electrode over Fp2. In this study, there were two experimental groups, which differed on the modality of feedback presentation: in the three dimensions (3D) virtual reality group, participants received the feedback while they were in the middle of a virtual room, whereas the two dimensions (2D) virtual reality group simply watched a cinema screen. The authors predicted that the 3D virtual reality modality was more effective, based on Gruzelier et al. (2010) findings of faster learning rates when feedback is delivered in a 3D virtual reality environment. In line with this hypothesis, they found that only the 3D virtual reality group achieved a significant electrophysiological learning across sessions, in terms of learning scores, calculated as the points awarded for exceeding

TABLE 9 Statistics and results of theta neurofeedback protocols with older adults

		EEG			Behavioral		
	Measure	Training × group interaction	Training effect on NFT group	Measure	Training × group interaction	Training effect on NFT group	NFT success
Reis et al. (2016)	ind_${}_{\text{fm}}$Theta, baseline	NA (NT)	**Wilcoxon signed-rank, $p = .037$**	Digit Span, ACC	NA (NT)	Wilcoxon signed-rank, $p = .219$	Not Successful
	ind_${}_{\text{fm}}$Theta, activity	NA (NT)	**Wilcoxon signed-rank, $p = .010$**				
Wang and Hsieh (2013)	${}_{\text{fm}}$Theta	**${}_{\text{mixed}}F = 10.39$, *$p = .003$, $d = 1.14$***	**${}_{\text{2t\_paired}}t = 3.43$, *$p = .008$, $d = 1.21$***	ANT_Conflict, RT	**${}_{\text{mixed}}F = 32.59$, *$p < .001$, $d = 2.02$***	${}_{\text{2t\_paired}}t = 8.17$, $p < .001$, $d = 2.58$	**Successful**
				ANT_Conflict, ACC	NS (NR)	NA (NT)	
				WM	**${}_{\text{mixed}}F = \text{NR}$, $p < .05$, d = NR**	${}_{\text{2t\_paired}}t = 3.21$, *$p = .006$, $d = 2.58$*	

Note: Bold text indicates significant results; italic text indicates results that were not reported in the original article but could be estimated from available data; underlined text indicates reported results that we transformed for the sake of homogeneity.

Abbreviations: ANT, attention network task; ${}_{\text{fm}}$Theta, frontal-midline theta; ind_${}_{\text{fm}}$Theta, individual frontal-midline theta; NA, not available; NFT, neurofeedback training; NR, not reported; NS, not significant; NT, not tested.

the threshold levels of frontal alpha power. Specifically, they found a significant interaction between Training phase and Group, which was qualified by a significant increase in alpha learning scores in the 3D NFT group. Behavioral performance was assessed using the Stroop task and calculating the Gratton effect, according to which Stroop effect on trial n is reduced if the preceding trial n-1 was incongruent compared to a congruent one. They found that NFT reduced the Gratton effect as shown by a performance improvement on trials succeeding a congruent compared to an incongruent trial. This occurred for both accuracy and RTs in the 3D group and only for RTs in the 2D group, suggesting that NFT enhanced response inhibition and conflict monitoring, and this improvement was greater when NFT was delivered in a 3D environment. Interestingly, both for electrophysiological and behavioral results, the effect sizes were quite large. Taken all together, these results need to be interpreted with caution since they exclusively suggest that 3D NFT is more effective than 2D NFT. Indeed, although this study provided evidence for NFT success both at electrophysiological and behavioral levels, it does not fully satisfy our research question, which could have been resolved only by comparing the NFT group to a control group not receiving real NFT at all.

Concerning unsuccessful studies, Escolano and colleagues (2011) focused on upper alpha frequency averaged over parieto-occipital locations (electrodes: P3, Pz, P4, O1, and O2), using a single-band protocol and training participants for a total of 125 min of NFT. Electrophysiological results showed changes in terms of linear increase in upper alpha both during a counting task, defined as “active” measurement and during passive resting state EEG. Specifically, participants receiving NFT showed a significant difference in power between pre-active assessment of session 5 and the analog block in the first session. Great caution must be taken when considering these electrophysiological effects, because they refer to the experimental group only as EEG was not measured for the control group, who performed solely the working memory task at the beginning and at the end of the study. Moreover, another major drawback is that 3 out of 10 NFT group participants were qualified as nonresponders and were excluded from analyses, which included only responder participants. Behavioral analyses suffer from a pitfall as well, as the authors claimed that NFT participants enhanced their working memory performance on the conceptual span task only by comparing pretraining and posttraining scores. Hence, in contrast with authors' conclusions, the fact that no between-groups comparisons were reported precluded solid evidence in favor of this protocol efficacy.

In a subsequent study, Escolano et al. (2014) overcame previous limitations, designing a single session upper alpha NFT study (in total 25 min) and including a sham control group. At the electrophysiological level, the Training phase by Group interaction did not reach significance. Therefore,

TABLE 10 Statistics and results of alpha neurofeedback protocols with younger adults

		EEG			Behavioral		NFT success
	Measure	Training × group interaction	Training effect on NFT group	Measure	Training × group interaction	Training effect on NFT group	
Berger and Davelaar (2018)	$_{\text{fp}}$Alpha learning scores	**$_{\text{mixed}}F = 7.97$, $p = .010$, $d = 1.20$**	**$F = 4.35$, $p = .042$, $R^2 = 0.06$**	Gratton Effect, ACC	**$_{\text{mixed}}F = 5.17$, *$p = .035$*, $d = 0.97$**	**$_{\text{2t\_paired}}\ t = 2.33$, *$p = .045$*, $d = 0.70$**	**Successful**
				Gratton Effect, RT	NS (NR)	NA (NT)	
Escolano et al. (2011)	$_{\text{po}}$Upper Alpha, resting state	NA (NT)	**$_{\text{2t\_paired}}t = 3.79$, $p = .006$, *$d = 1.55$***	Conceptual Span, ACC	NA (NT)	**$_{\text{2t\_paired}}\ t = 2.34$, *$p = .044$*, *$d = 0.74$***	Not Successful
	$_{\text{po}}$Upper Alpha, active tasks	NA (NT)	**$_{\text{2t\_paired}}t = 3.87$, $p = .006$, *$d = 1.58$***				
Escolano et al. (2014)	$_{\text{po}}$Upper Alpha, resting state	$_{\text{mixed}}\ F = 3.84$, $p = .801$, *$d = 0.90$*	**$_{\text{2t\_paired}}t = 3.97$, *$p = .003$*, *$d = 1.26$***	TMT-B	**$_{\text{mixed}}F = 4.51$, $p = .049$, *$d = 0.97$***	**$_{\text{2t\_paired}}t = 4.26$, $p = .002$, *$d = 1.35$***	Partially successful
				PASAT, ACC	NS (NR)	**$_{\text{2t\_paired}}t = 3.05$, $p = .014$, *$d = 0.96$***	
				PASAT, RT	NS (NR)	**$_{\text{2t\_paired}}t = 5.28$, $p < .001$, *$d = 1.67$***	
				RAVLT, ACC	NS (NR)	**$_{\text{2t\_paired}}t = 2.59$, $p = .029$, *$d = 0.82$***	
Gordon et al. (2020)	ind_$_{\text{p}}$Upper Alpha, NFT versus silent control	$_{\text{mixed}}F = 0.07$, $p = .801$, $d = 0.07$	NA (NT)	WM	$_{\text{mixed}}\ F = 0.102$, $p = .749$, *$d = 0.10$*	NA (NT)	Not successful
	ind_$_{\text{p}}$Upper Alpha, NFT versus active control	$_{\text{mixed}}F = 0.30$, $p = .583$, $d = 0.14$	NA (NT)				
Hsueh et al. (2016)	Alpha	**$_{\text{mixed}}F = 7.05$, $p < .001$, *$d = 0.75$***	**$_{\text{2t\_paired}}t = \text{NR}$, $p < .05$, $d = \text{NR}$**	Backward Digit Span, ACC	$_{\text{mixed}}F = 0.43$, $p = .516$, *$d = 0.19$*	$_{\text{2t\_paired}}t = 0.66$, $p = .516$, *$d = 0.13$*	Partially successful
				Operation Span, ACC	$_{\text{mixed}}F = 1.13$, $p = .294$, *$d = 0.30$*	$_{\text{2t\_paired}}t = 1.06$, $p = .294$, *$d = 0.21$*	
Naas et al. (2019)	ind_$_{\text{po}}$Upper Alpha	$_{\text{mixed}}F = 0.58$, $p = .363$, $d = 0.27$	**$_{\text{2t\_paired}}t = 2.63$, $p = .018$, *$d = 0.64$***	Digit Span, ACC	$_{\text{mixed}}F = 1.24$, $p = .280$, $d = 0.39$	$_{\text{2t\_paired}}t = 1.85$, $p = .083$, *$d = 0.45$*	Not successful
Pei et al. (2018)	Alpha	NA (NT)	**$_{\text{2t\_paired}}t = 2.91$, $p = .017$, *$d = 0.92$***	Backward Digit Span, ACC	NA (NT)	**$_{\text{2t\_paired}}t = 4.28$, $p = .002$, *$d = 1.35$***	Not successful
Wei et al. (2017)	Alpha power ratio	**$_{\text{mixed}}F = 8.35$, $p < .001$, *$d = 1.06$***	**$_{\text{2t\_paired}}t = \text{NR}$, $p < .05$, $d = \text{NR}$**	Backward Digit Span, ACC	$_{\text{mixed}}F = 3.31$, $p = .079$, *$d = 0.66$*	**$_{\text{2t\_paired}}t = \text{NR}$, $p < .05$, $d = \text{NR}$**	Partially successful

Note: Bold text indicates significant results; italic text indicates results that were not reported in the original article but could be estimated from available data; underlined text indicates reported results that we transformed for the sake of homogeneity.

Abbreviations: $_{\text{fp}}$Alpha, fronto-parietal alpha; ind_$_{\text{p}}$Upper Alpha, individual parietal upper alpha; ind_$_{\text{po}}$Upper Alpha, individual parieto-occipital upper alpha; NA, not available; NFT, neurofeedback training; NR, not reported; NS, not significant; NT, not tested; PASAT, paced auditory serial addition task; $_{\text{po}}$Upper Alpha, parieto-occipital upper alpha; RAVLT, Rey auditory verbal learning test; TMT-B, trail making test part B; WM, working memory.

the authors' claim that upper alpha was significantly enhanced after a short NFT period in the NFT group only should be taken with caution, as no differential effect was reported. Concerning NFT behavioral effect, this study provided initial evidence for executive function enhancement, as documented by the significant interaction between Training and Group for the part B of the Trail Making Test, which was qualified by a positive progress after the NFT session. Taken as a whole, we classified this study as partially successful, since electrophysiological results do not allow to clearly demonstrate that NFT induced significant changes in the experimental group compared to the control one.

Proceeding with “not successful” studies, Gordon and colleagues (2020) required participants to upregulate parietal individual upper alpha measured from Pz electrode for a total time of 150 min. Interestingly, participants were divided into six groups: two combined groups (NFT + WMT group, in which NFT was delivered along with a working memory training and NFT + active control training, namely, a visual search training), three single-protocol groups (NFT, WMT, and active control training), and a passive control group. To inquire NFT influence on upper alpha power, they contrasted NFT only group with active and passive control groups, but they did not find any Training phase by Group interaction on resting state EEG. The effects on behavioral performance were analyzed comparing NFT + WMT and WMT-only groups and revealed no significant difference, suggesting that NFT did not yield working memory enhancements. Of note, the feedback was provided using an innovative modality, that is, participants did not receive a classical visual feedback but were rewarded with points which allowed them to progress in the game they were playing. Thus, this procedure might have limited the NFT behavioral effects.

Naas et al. (2019) targeted individual upper alpha over parietal and occipital regions (P7, O1, O2, and P8) for 60 min in total. Although they reported a significant increase in individual upper alpha from period 1 to period 20 in the NFT group, the Training phase by Group interaction did not reach significance, indicating that there was no electrophysiological difference between the experimental group and the sham control group. Moreover, the posttraining alpha level was significantly correlated with participants' pretraining alpha level, suggesting that the initial alpha power was the best predictor of individual upper alpha improvement. Similarly, this protocol did not produce any change in working memory on the eight digit-span test, as suggested by the absence of a significant Training phase by Group interaction.

Pei and colleagues (2018) trained alpha band measured from Fz and C4 electrodes during a five-session experiment (in total 180 min of NFT) and compared the experimental NFT group with a sham-NFT group. To investigate across group differences, they contrasted the alpha power levels reached by the two groups in the last session. However, since this analysis does not allow to highlight a true Training phase by Group interaction, we classified electrophysiological effects as not significant. Behavioral analyses were conducted likewise, by comparing the two groups' posttraining accuracies on the backward digit span task. Therefore, due to the absence of a direct statistical comparison between groups that would also take into account baseline measurement, this protocol did not provide convincing evidence for NFT efficacy on working memory.

The two remaining studies were classified as “partially successful” as they were both effective solely at the electrophysiological level. Hsueh et al. (2016) designed an alpha NFT protocol of 12 sessions for a total of 432 min of training. They applied a bipolar montage over central regions by locating 6 electrodes at 2.5 cm anteriorly and posteriorly to C3, Cz, and C4, respectively, and used a sham random frequency control group. They found a significant interaction between Training phase and Group for mean alpha amplitude, suggesting that, compared to earlier sessions and to the sham group, mean alpha amplitude was higher during the last sessions (8th to 12th) in the NFT group. Behavioral effects were investigated using the backward digit span task, that is, requiring participants to reverse the order of the previously seen digits, and the operation span task, during which they were asked to recall three letters presented in the learning trials, but, in between, they were required to judge the accuracy of an intervening mathematical equation. For neither of them the interaction between the factors Training phase and Group was significant, revealing no NFT effect on working memory performance.

Lastly, Wei and colleagues (2017) trained participants to upregulate alpha over C3 during 12 sessions for a total of 300 min. The analyses of electrophysiological effects revealed that the interaction between Training and Group was significant, suggesting that, compared to sham control group, participants receiving alpha NFT had a progressive alpha power increase throughout the sessions. These results were not accompanied by equally successful outcomes for working memory performance, as shown by the absence of a significant Training phase by Group interaction concerning the backwards digit span.

3.4 | Alpha protocols with older adults

Alpha NFT with older adults was among the least effective protocols as it obtained a SR of 0% (Table 11). Reis and colleagues' (2016) study has already been presented in the theta protocol section, as it targeted at the same time and in the same direction two frequencies, that is, alpha and theta. Since NFT participants received at the same time alpha and theta training, their effects cannot be disentangled and, thus, it is not clear whether one of the two frequency bands was

TABLE 11 Statistics and results of alpha neurofeedback protocols with older adults

	EEG			Behavioral			
	Measure	**Training × group interaction**	**Training effect on NFT group**	**Measure**	**Training × group interaction**	**Training effect on NFT group**	**NFT success**
Gomez-Pilar et al. (2016)	Upper Alpha	NA (NT)	Wilcoxon signed rank, *p* = .235	ANT_Conflict	Mann-Whitney, *p* = .986	Wilcoxon signed rank, *p* = .137	Not successful
Reis et al. (2016)	ind_Alpha, baseline	NA (NT)	**Wilcoxon signed-rank, *p* = .049**	Digit span, ACC	NA (NT)	Wilcoxon signed-rank, *p* = .219	Not successful
	ind_Alpha, activity	NA (NT)	NS (NR)				

Note: Bold text indicates significant results; italic text indicates results that were not reported in the original article but could be estimated from available data; underlined text indicates reported results that we transformed for the sake of homogeneity.

Abbreviations: ANT, attention network task; ind_Alpha, individual alpha; NA, not available; NFT, neurofeedback training; NR, not reported; NS, not significant; NT, not tested.

predominantly responsible for the lack of success. In line with theta NFT effects, this protocol was "not successful" as no interaction effect was reported to directly compare NFT group to controls. Therefore, despite the increase in alpha power during baseline after NFT, we categorized electrophysiological effects as not significant. Moreover, NFT did not yield significant enhancements on tasks tapping on working memory (i.e., digit span task).

Similarly, Gomez-Pilar et al. (2016) study was "not successful". In this case, the alpha NFT protocol required participants to suppress this frequency while upregulating beta over central regions (training electrodes: C3, Cz, and C4). Despite the high amount of NFT (450 min), this protocol did not produce the expected effects on alpha frequency, which was unaltered in the posttraining assessment. Moreover, posttraining executive function performance, measured with attentional control subtest contained in the Luria-AND battery, did not show enhancements on tasks tapping on response inhibition, during which participants were required to inhibit automatic responses and select less habitual ones.

3.5 | Low beta protocols with younger adults

Low beta NFT protocols with younger adults achieved a SR of 0%, with three studies "not successful" and four "partially successful" (Table 12).

Starting with the "not successful" studies, Cannon et al. (2009) implemented a low-resolution electromagnetic tomographic (LORETA) NFT to train frequency-specific activity at the cortical level. Thus, instead of targeting electrical activity from scalp electrodes, as the other studies here reviewed, this study targeted specifically cortical electrical activation. More in detail, participants were trained to upregulate low beta (14–18 Hz) within one of the following cortical regions of training (ROTs): anterior cingulate cortex (ACC) and left and right dorsolateral prefrontal cortices (lDLPFC and rDLPFC). Thus, three experimental groups were compared. Participants received a total of 528 min of NFT, for 33 sessions. EEG measures throughout the sessions reported a significant learning effect in low beta band, found specifically in the ACC group. This effect, however, did not provide compelling evidence for electrophysiological effects as it was not substantiated by any Training phase by Group interaction. Moreover, the authors declared that ACC participants were the only ones to show significant learning scores in posttraining behavioral assessment. More specifically, after NFT, ACC group produced higher Working Memory Index (WMI) scores, estimated from WAIS–III subtests (arithmetic, digit span, and letter-number sequencing). However, such enhancements in working memory could not be inferred from the data reported in the article, which were limited to a figure presenting WMI after training. Hence, due to the lack of

TABLE 12 Statistics and results of low beta neurofeedback protocols with younger adults

	EEG			Behavioral			
	Measure	**Training × group interaction**	**Training effect on NFT group**	**Measure**	**Training × group interaction**	**Training effect on NFT group**	**NFT success**
Cannon et al. (2009)	Beta in ACC	NA (NT)	**$F = 141.40$, $p < .001$,** ***$d = 4.20$***	WMI	NA (NT)	NA (NT)	Not successful
Egner and Gruzelier (2004) (beta 1)	NA (NT)	NA (NT)	NA (NT)	Go/Nogo, RT	NS (NR)	**$_{\text{2t\_paired}}t = 3.65$,** ***$p = .008$, $d = 1.29$***	Not successful
				Go/Nogo, FA	NS (NR)	NA (NT)	
				Go/Nogo, d’	NS (NR)	NA (NT)	
Egner and Gruzelier (2004) (SMR)	NA (NT)	NA (NT)	NA (NT)	Go/Nogo, RT	NS (NR)	NS (NR)	Not successful
				Go/Nogo, FA	NS (NR)	NA (NT)	
				Go/Nogo, d’	NS (NR)	$_{\text{2t\_paired}}t = 1.75$, $p = .058$, *$d = 0.58$*	
Gonçalves et al. (2018)	SMR	**$_{\text{mixed}}F = 6.44$, $p = .017$,** ***$d = 0.93$***	**$_{\text{2t\_paired}}t = 4.99$, $p < .001$,** ***$d = 1.29$***	ANT_Conflict, RT	NA (NT)	$_{\text{2t\_paired}}t = 2.00$, $p = .064$, *$d = 0.52$*	Partially successful
Kober et al. (2017)	SMR	**$_{\text{mixed}}F = 6.18$, $p < .05$**	NA (NT)	CBTT Backwards, ACC	NA (NT)	**$_{\text{2t\_paired}}t = 2.79$, $p = .021$,** ***$d = 0.88$***	Partially successful
				Digit Span Backwards, ACC	NA (NT)	**$_{\text{2t\_paired}}t = 2.75$, $p = .022$,** ***$d = 0.87$***	
Vasquez et al. (2015)	$_{\text{fl}}$Beta	NA (NR)	NA (NT)	IOWA	**$_{\text{mixed}}F = 3.68$, $p = .039$,** ***$d = 0.70$***	**$_{\text{2t\_paired}}t = 2.52$, $p = .033$,** ***$d = 0.80$***	Partially successful
Vernon et al. (2003)	SMR/Theta ratio	NA (NT)	**$_{\text{2t\_paired}}t = 3.15$, $p = .012$,** ***$d = 1.00$***	Conceptual Span, ACC	**$_{\text{mixed}}F = 5.20$, $p = .004$,** ***$d = 0.83$***	**$_{\text{2t\_paired}}\underline{t = 5.14}$, $\underline{p < .001}$,** ***$d = 1.62$***	Partially successful
	SMR/Beta ratio	NA (NT)	$_{\text{2t\_paired}}t = 2.02$, $p = .074$, *$d = 0.64$*				

Note: Bold text indicates significant results; italic text indicates results that were not reported in the original article but could be estimated from available data; underlined text indicates reported results that we transformed for the sake of homogeneity.

Abbreviatons: ANT, attention network task; CBTT, Corsi block tapping test; $_{\text{fl}}$Beta, fronto-lateral beta; NA, not available; NS, not significant; NFT, neurofeedback training; NR, not reported; NT, not tested.

precise statistics supporting behavioral results, we could not consider them as significant.

Targeting low beta in the 15- to 18-Hz frequency range measured from Cz, Egner and Gruzelier (2004) trained participants to upregulate this frequency for 150 min, comparing the experimental group with a control group that received a behavioral training of equal duration of NFT and with another experimental group receiving NFT for upregulating SMR. Executive function changes were measured on a visual Go/Nogo task in which participants were instructed to respond as quickly as possible to targets and to refrain responses to nontargets. NFT effects in the EEG domain were not verified by directly investigating the trained frequency band but targeted P300 amplitude at C3, Cz, and Pz, which was considered as an index of integration of task-relevant information in working memory. Although they found that beta1 NFT induced a significant increase of target P300 amplitude, this result cannot be regarded as direct evidence that NFT induced changes in the trained frequency band. An additional weakness was that EEG measures were not collected for the control group, and the beta1 group was not compared with the SMR group. Consequently, evidence for specific electrophysiological effects of beta1 NFT was hampered by the lack of comparisons between groups. We should sound a note of caution with regard to behavioral results as well. In fact, despite a reported response inhibition enhancement, in terms of reduced RTs on the Go/Nogo task compared to the control group, the interaction between Training phase and Group was not significant. As previously mentioned, Egner and Gruzelier' s (2004) study comprised a second experimental protocol, namely SMR NFT, which was “not successful” either. Likewise, it suffered from the same pitfalls and SMR NFT did not produce either posttraining EEG changes or executive function enhancements.

The remaining studies belong to the “partially successful” category. Gonçalves and colleagues (2018), in addition to the theta NFT protocol, implemented a SMR NFT during which participants were asked to upregulate SMR and downregulate theta, namely, the opposite of the protocol discussed above (see younger adults theta protocol section). The SMR protocol was effective at increasing SMR amplitude as indicated by the significant Training phase by Group interaction. In addition, paired-samples *t*-tests revealed larger SMR amplitude when comparing baseline to the last block. Despite these EEG changes, no conflict monitoring improvement was observed in terms of conflict scores on the ANT. The short duration of NFT (25 min) might have prevented executive function improvements.

Similar outcomes were obtained by Kober and coworkers (2017), who implemented NFT to upregulate SMR measured from Cz for a total time of 450 min and compared it to another experimental group receiving a gamma upregulation protocol. Working memory improvements were assessed using the Corsi Block Tapping test (CBTT) backwards task and the digit span backwards task. During the former, subjects were asked to tap on the blocks previously tapped by the experimenter but in the reverse order, whereas the latter requires participants to recall a sequence of digits backwards. At the end of NFT, SMR power, measured during resting state EEG, showed a significant interaction between Training phase and Group, indicating that the SMR group increased SMR compared to pretraining and to the gamma group. By contrast, working memory performance was not compared with that in the other experimental group, not allowing firm conclusions about possible behavioral improvements.

Vasquez and collaborators (2015) required participants to upregulate beta (13–21 Hz) while downregulating theta for a single NFT-session of 30 min. Inhibition performance was enhanced on the Iowa Gambling Task but solely for the right hemisphere group (as discussed in theta section). Indeed, the significant interaction between Training and Group was explained by the significant statistical changes in the posttraining phase in the right hemisphere group only. However, no electrophysiological changes in beta were found.

Lastly, Vernon and colleagues (2003) provided participants with 120 min of multiband NFT protocol to upregulate SMR while downregulating theta and beta measured from Cz. They compared two experimental groups, namely, SMR and theta. We have already discussed the latter in the theta section (theta upregulation, delta, and alpha downregulation), showing that it was “not successful”. Since they did not report any Training phase by Group interaction, the increase in SMR/theta ratio could not be regarded as evidence for NFT electrophysiological effects. By contrast, behavioral outcomes of SMR training were clearer as there was a significant interaction between Training phase and Group, suggesting that, compared with the theta group, the SMR group showed significant working memory enhancements on a conceptual span task, with higher accuracy in the posttraining assessment.

3.6 | Low beta protocols with older adults

Older adults low beta protocols achieved a SR of 0%: one study was “partially successful” and the other “not successful” (Table 13).

The “partially successful” study was conducted by Campos da Paz and collaborators (2018) and targeted SMR. Older participants were trained for 90 min to upregulate SMR in central regions comparing their performance with a sham control group, which received NFT only for the first training session, and with a passive control group. NFT participants showed pre-post training changes at the electrophysiological level, in terms of decreased activation in all frequency bands compared to the sham control group, which showed an increased activation. The authors suggested that the lower activation of the NFT group reflected more efficient cortical integration, whereas the sham control group higher activation was interpreted as compensatory activity. However, since they did not test NFT effects on the trained frequency band, we could not draw conclusions about NFT electrophysiological specificity. By contrast, they found a significant Training phase by Group interaction on the delayed matched to sample task, that is, a visual working memory task in which participants viewed a stimulus and, after an interval, had to point to the first stimulus, choosing between the previous one and the new one. Moreover, the NFT group showed higher accuracy after the training.

In the protocol implemented by Gomez-Pilar and colleagues (2016), participants were required to upregulate beta (18–21 Hz) while suppressing alpha (see older adults alpha protocol section). EEG changes in the intended direction were observed specifically for beta, both for 18 and 21 Hz but not for alpha. Nonetheless, no group comparison was performed as EEG was not measured in the control group. Moreover, no executive function improvement was found on response inhibition, measured with the attentional control subtest of the Luria-AND battery.

TABLE 13 Statistics and results of low beta neurofeedback protocols with older adults

		EEG			Behavioral		
	Measure	**Training × group interaction**	**Training effect on NFT group**	**Measure**	**Training × group interaction**	**Training effect on NFT group**	**NFT success**
Campos da Paz et al. (2018)	SMR	NA (NT)	NA (NT)	Delayed Match to Sample, ACC	**$_{\text{mixed}}F = 10.25$, *$p < .001$*, *$d > 1.00$***	***$_{2t\_paired}t = 5.02$*, $p = .002$, *$d = 1.90$***	Partially Successful
Gomez-Pilar et al. (2016)	Beta (centered in 18 Hz)	NA (NT)	**Wilcoxon signed rank, $p < .001$**	ANT_Conflict	Mann-Whitney, $p = .986$	Wilcoxon signed rank, $p = .137$	Not Successful
	Beta (centered in 21 Hz)	NA (NT)	**Wilcoxon signed rank, $p = .002$**				

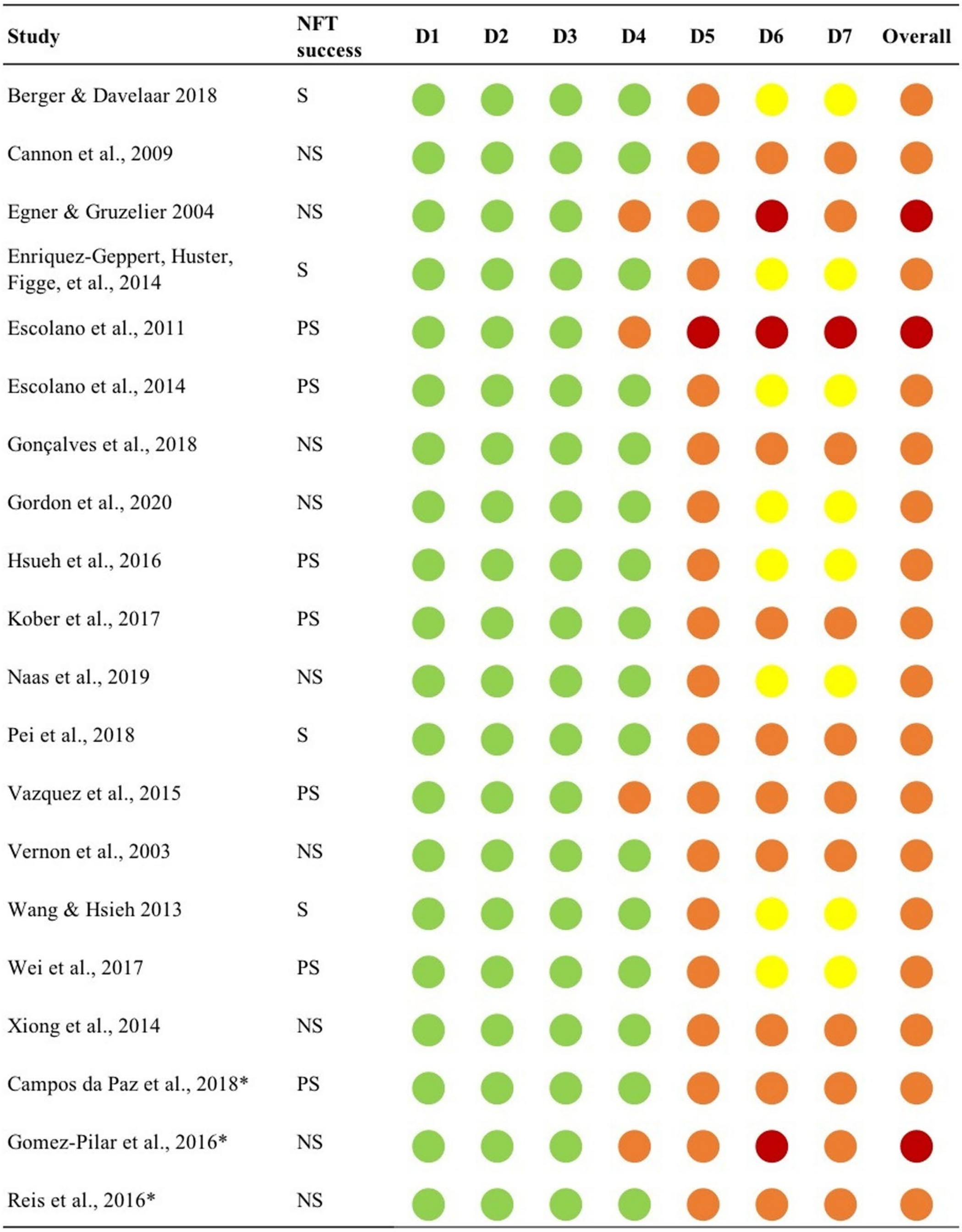
Note: Bold text indicates significant results; italic text indicates results that were not reported in the original article but could be estimated from available data; underlined text indicates reported results that we transformed for the sake of homogeneity.

Abbreviations: ANT, attention network task; NA, not available; NFT, neurofeedback training; NR, not reported; NS, not significant; NT, not tested.

3.7 | Risk of bias and quality assessment

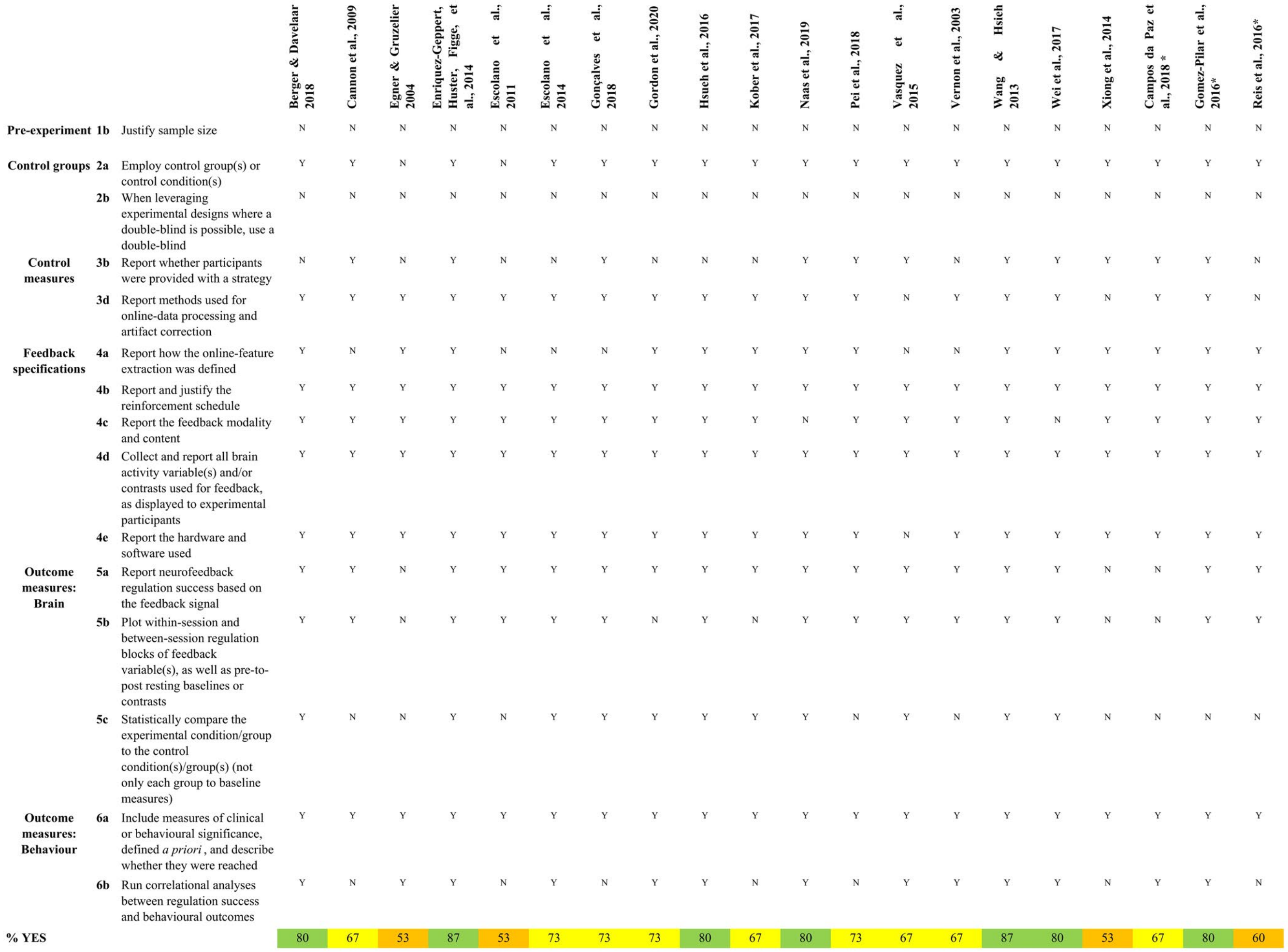
The most striking result that emerged from the qualitative analysis of the three NFT protocols is that many of the studies suffered from methodological drawbacks. Indeed, the employed statistical approach was often not appropriate for investigating whether the experimental group exhibited electrophysiological and behavioral changes following NFT and compared to the control group. Moreover, the lack of shared and robust experimental designs for NFT studies might have negatively affected the outcomes. However, it is critical to note that effect sizes were generally from medium to large or very large. We interpret this finding with caution, as we suspect that effect sizes might be inflated, probably due to the underpowered nature of many studies and publication bias.

For all these reasons, and to have a broader view of the quality of the analyzed evidence, we decided to include two additional measures to evaluate the studies included in the present review: the Risk Of Bias In Nonrandomized Studies–Of Interventions (ROBINS–I, Sterne et al., 2016) and the Consensus on the reporting and experimental design of clinical and cognitive-behavioral neurofeedback studies (CRED–nf checklist, Ros et al., 2020). The first is a tool for the assessment of risk of bias in estimates of the effectiveness of an intervention when studies are not randomized, which was the case of all our included studies. It comprises seven domains through which bias might be introduced in a nonrandomized study and whose judgment can vary between low, moderate, serious, critical risk of bias, or no information. Each domain contributes to the overall risk of bias judgment (see Figure 2). On the other hand, the CRED–nf checklist is a recent tool specific to NFT works, which stemmed from the need of designing studies with methodological rigour, considered as the only way to advance the field of NFT and to better highlight its underlying mechanisms. For our purposes, we evaluated only the essential items using a binary approach (yes or no) to obtain an overall percentage level indicating to what extent the included studies satisfied the criteria required for well-designed NFT protocols (see Figure 3).



Study	NFT success	D1	D2	D3	D4	D5	D6	D7	Overall
Berger & Davelaar 2018	S	🟢	🟢	🟢	🟢	🟠	🟡	🟡	🟠
Cannon et al., 2009	NS	🟢	🟢	🟢	🟢	🟠	🟠	🟠	🟠
Egner & Gruzelier 2004	NS	🟢	🟢	🟢	🟠	🟠	🔴	🟠	🔴
Enriquez-Geppert, Huster, Figge, et al., 2014	S	🟢	🟢	🟢	🟢	🟠	🟡	🟡	🟠
Escolano et al., 2011	PS	🟢	🟢	🟢	🟠	🔴	🔴	🔴	🔴
Escolano et al., 2014	PS	🟢	🟢	🟢	🟢	🟠	🟡	🟡	🟠
Gonçalves et al., 2018	NS	🟢	🟢	🟢	🟢	🟠	🟠	🟠	🟠
Gordon et al., 2020	NS	🟢	🟢	🟢	🟢	🟠	🟡	🟡	🟠
Hsueh et al., 2016	PS	🟢	🟢	🟢	🟢	🟠	🟡	🟡	🟠
Kober et al., 2017	PS	🟢	🟢	🟢	🟢	🟠	🟠	🟠	🟠
Naas et al., 2019	NS	🟢	🟢	🟢	🟢	🟠	🟡	🟡	🟠
Pei et al., 2018	S	🟢	🟢	🟢	🟢	🟠	🟠	🟠	🟠
Vazquez et al., 2015	PS	🟢	🟢	🟢	🟠	🟠	🟠	🟠	🟠
Vernon et al., 2003	NS	🟢	🟢	🟢	🟢	🟠	🟠	🟠	🟠
Wang & Hsieh 2013	S	🟢	🟢	🟢	🟢	🟠	🟡	🟡	🟠
Wei et al., 2017	PS	🟢	🟢	🟢	🟢	🟠	🟡	🟡	🟠
Xiong et al., 2014	NS	🟢	🟢	🟢	🟢	🟠	🟠	🟠	🟠
Campos da Paz et al., 2018*	PS	🟢	🟢	🟢	🟢	🟠	🟠	🟠	🟠
Gomez-Pilar et al., 2016*	NS	🟢	🟢	🟢	🟠	🟠	🔴	🟠	🔴
Reis et al., 2016*	NS	🟢	🟢	🟢	🟢	🟠	🟠	🟠	🟠

FIGURE 2 Risk of bias assessment with ROBINS–I (Sterne et al., 2016). The risk ok bias of each study in the seven domains was evaluated and then the overall risk of bias judgment was formulated. Notes: D1, bias due to confounding; D2, bias in selection of participants into the study; D3, bias in classification of interventions; D4, bias due to deviations from intended interventions; D5, bias due to missing data; D6, bias in measurement of outcomes; D7, bias in selections of the reported results; green, low risk of bias; yellow, moderate risk of bias; orange, serious risk of bias; red, critical risk of bias. * = Studies with older adults only



			Berger & Davelaar 2018	Cannon et al., 2009	Egner & Gruzelier 2004	Enriquez-Geppert, Huster, Figge, et al., 2014	Escolano et al., 2011	Escolano et al., 2014	Gonçalves et al., 2018	Gordon et al., 2020	Hsueh et al., 2016	Kober et al., 2017	Naas et al., 2019	Pei et al., 2018	Vasquez et al., 2015	Vernon et al., 2003	Wang & Hsieh 2013	Wei et al., 2017	Xiong et al., 2014	Campos da Paz et al., 2018 *	Gomez-Pilar et al., 2016*	Reis et al., 2016*
Pre-experiment	**1b**	Justify sample size	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
Control groups	**2a**	Employ control group(s) or control condition(s)	Y	Y	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	2b	When leveraging experimental designs where a double-blind is possible, use a double-blind	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
Control measures	**3b**	Report whether participants were provided with a strategy	N	Y	N	Y	N	N	Y	N	N	N	Y	Y	Y	N	Y	Y	Y	Y	Y	N
	3d	Report methods used for online-data processing and artifact correction	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	N	Y	Y	N
Feedback specifications	**4a**	Report how the online-feature extraction was defined	Y	N	Y	Y	N	N	N	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Y
	4b	Report and justify the reinforcement schedule	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	4c	Report the feedback modality and content	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Y	Y	Y	Y
	4d	Collect and report all brain activity variable(s) and/or contrasts used for feedback, as displayed to experimental participants	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	4e	Report the hardware and software used	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y
Outcome measures: Brain	**5a**	Report neurofeedback regulation success based on the feedback signal	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	Y	Y
	5b	Plot within-session and between-session regulation blocks of feedback variable(s), as well as pre-to-post resting baselines or contrasts	Y	Y	N	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Y	Y	N	N	Y	Y
	5c	Statistically compare the experimental condition/group to the control condition(s)/group(s) (not only each group to baseline measures)	Y	N	N	Y	N	Y	Y	Y	Y	Y	Y	N	Y	N	Y	Y	N	N	N	N
Outcome measures: Behaviour	**6a**	Include measures of clinical or behavioural significance, defined *a priori* , and describe whether they were reached	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	6b	Run correlational analyses between regulation success and behavioural outcomes	Y	N	Y	Y	N	Y	N	Y	Y	N	Y	N	Y	Y	Y	Y	N	Y	Y	N
% YES			80	67	53	87	53	73	73	73	80	67	80	73	67	67	87	80	53	67	80	60

FIGURE 3 Assessment of neurofeedback training (NFT) protocol quality with CRED–nf checklist (Ros et al., 2020). We assessed whether the included studies satisfied the criteria for well-designed NFT protocols by answering with "yes" or "no" to each essential item. Y, yes; N, no; *, studies including older adults only

4 | DISCUSSION

Our systematic review of EEG-NFT literature sought to address whether existing NFT studies have provided convincing evidence for this technique effectiveness at enhancing executive functions. We found that EEG-NFT with healthy individuals is attracting considerable interest, but only a small extent of studies has specifically targeted executive functions. Therefore, our qualitative analysis of the literature was conducted on twenty works, some of which comprised more than a single NFT protocol.

Overall, to date, the evidence for NFT efficacy in enhancing healthy adults' executive functions is flimsy, as indicated by a SR of 14.29% when considering all the retrieved NFT protocols with younger adults and 20% for all the included NFT protocols with older adults. Crucially, we used a strict criterion to define the study success, as it required the achievement of significant modulations both in the EEG and in the executive function domains. In fact, when calculating the SR, we did not include partially successful studies that obtained significant results only at one level, and, to define a study as "successful", we assessed whether it satisfied both trainability and interpretability criteria (Zoefel et al., 2011). Moreover, our judgment was based only on the significance of the interaction between Training phase and Group, that is, when the NFT and the control groups were directly compared to each other. In fact, this represents the only reasonable way to attest whether, after the training, the group receiving NFT outperformed the control group on the target electrophysiological and behavioral measures (Nieuwenhuis et al., 2011).

The review was organized according to the frequency band targeted by NFT: theta, alpha, and low beta. This allowed us to address another fundamental aim, that is, to shed some light onto which of those targeted bands seems most effective and promising. As shown in Figure 4, younger adults' executive functions seemed to benefit more from theta NFT protocols (SR =33.33%), which, however, were the least frequent ones ($n = 6$), followed by alpha NFT, which were the most frequent protocols (we retrieved 8 out of 20 studies) and achieved a SR of 12.5%. On the other hand, none of the low

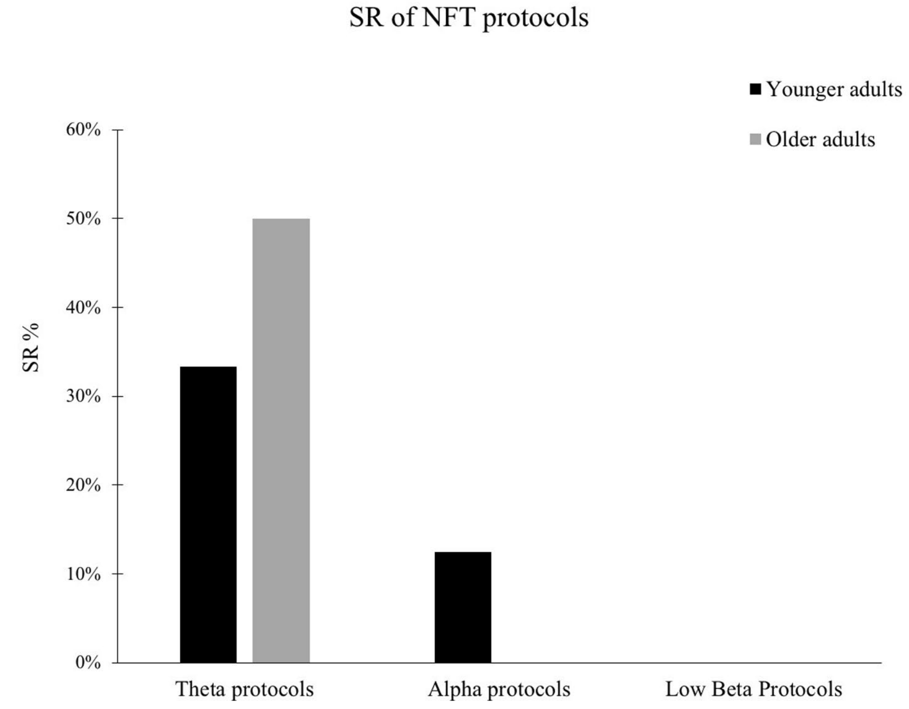


FIGURE 4 Comparison of the percentages of success ratio (SR) for the three neurofeedback training (NFT) protocols, divided into different ages (younger adults and older adults). Success ratio was calculated as the number of successful studies out of the total number of studies included in each protocol type

beta protocols ($n = 7$) was successful (SR = 0%). However, despite the differences in the SR, each NFT protocol showed its specificities, which we will now discuss.

Firstly, theta single-band protocols appeared to be more effective for enhancing executive functions and inducing the desired changes in the EEG domain, as shown by the fact that the successful experiments were those targeting solely theta frequency and specifically frontal-midline theta. However, in the discussed studies, the distinction between theta and frontal-midline theta NFT protocols was not fully clear. In fact, although solely Enriquez-Geppert, Huster, Figge and collaborators (2014) and Wang and Hsieh (2013) explicitly trained frontal-midline theta, whereas the other studies targeted more general theta, in practice, they all employed similar electrodes and an overlapping frequency range. Thus, in the analyzed studies, NFT protocols targeting frontal-midline theta cannot be clearly distinguished from those modulating theta in general. Moreover, other factors might explain the results of “partially successful” or “not successful” studies, such as the shorter duration of NFT in their protocols. In fact, the success of Enriquez-Geppert, Huster, Figge and collaborators (2014) and Wang and Hsieh (2013) might indicate that the amount of training needs to be high, at least more than 180 min, to induce significant NFT effects. Furthermore, our descriptive review pointed to the scarce efficacy of theta single-session NFT protocols (i.e., Gonçalves et al., 2018; Vasquez et al., 2015). Overall, our analysis suggested that NFT on frontal-midline theta measured from Fz (and neighboring electrodes) has been shown more frequently as effective in modulating this band and behavior than NFT protocols that did not include frontal-midline region electrodes. In fact, this type of protocol was shown to be effective for achieving improvements on memory updating and mental set-shifting (Enriquez-Geppert, Huster, Figge, et al., 2014) and on conflict monitoring (Wang & Hsieh, 2013) measures, with medium to very large effect sizes.

Alpha NFT protocols achieved a considerably lower success ratio and did not seem to be effective at training participants to increase this frequency band and at inducing executive function enhancements. Although NFT duration was high, generally more than 125 min, and single-band protocols were employed, only one study targeting this frequency in fronto-parietal scalp electrodes managed to increase alpha power and to enhance behavioral performance on response inhibition and conflict monitoring (Berger & Davelaar, 2018). However, as previously outlined, this study did not investigate specifically whether participants receiving NFT outperformed those who did not receive it at all. By contrast, the authors only provided evidence that 3D NFT induced greater executive function improvements than 2D NFT, in terms of Gratton effect reduction. It is interesting to note that a high number of alpha NFT protocols discussed here had methodological shortcomings, especially from the statistical point of view. Indeed, in most of these studies, the authors claimed to have obtained significant results which, however, could not be considered as reliable, direct evidence for NFT effectiveness. Such positive outcomes were achieved, for example, by limiting the analysis to responder participants. Escolano and colleagues (2011) excluded from the analysis of electrophysiological effects participants considered nonresponders, namely, those whose upper alpha power at the end of the training was not significantly higher than upper alpha at the beginning of the training. Similarly, Hsueh and coworkers (2016) did not find any Training phase by Group interaction at the level of behavioral performance, but their analysis on responders revealed working memory improvements after NFT. However, the procedure of excluding nonresponders seems to be risky, as it might bias results, favoring positive outcomes. Albeit the low SR of alpha NFT protocols, partially successful studies provided some promising pieces of evidence. Escolano and colleagues (2014) managed to induce working memory enhancements through parietal and occipital upper alpha single-session NFT, which, however, were not accompanied by electrophysiological training effects. Conversely, Hsueh and coworkers (2016) and Wei and colleagues (2017) modulated the alpha rhythm in the intended direction but without obtaining behavioral effects.

Lastly, low beta NFT success ratio was the lowest among younger adults' protocols. A distinction should be drawn as two specific subprotocols were included, namely, low beta NFT targeting a frequency range from 14 to 18 Hz and SMR (12–15 Hz), none of which was successful. Except for two single-session studies (Gonçalves et al., 2018; Vasquez et al., 2015), NFT duration was generally high (120–528 min) and this does not seem to be the cause. In contrast, we observed frequent methodological drawbacks, especially in the "not successful" studies. The major issue regarded the lack of direct comparisons between the NFT and the control groups. Specifically, many studies did not test interaction effects and limited their analysis to paired-sample *t*-tests, contrasting pretraining and posttraining measures in the experimental group. The "partially successful" studies suffer from limitations as well, but such weaknesses were usually limited to one level only, namely, either electrophysiological or behavioral. For example, Gonçalves and colleagues' (2018) study, despite its short duration, was well designed and it directly compared two different experimental groups, showing that SMR NFT induced a significant increase in SMR amplitude in the SMR group. However, in analyzing behavioral results, they did not test the interaction, hindering any insight into possible conflict monitoring enhancements. Likewise, Kober and coworkers (2017) found that SMR training was effective at the electrophysiological level, but their study was flawed by an inappropriate analysis for examining NFT effects on working memory performance. Hence, since partially unsuccessful outcomes could be explained more by the quality of the studies, in terms of lack of suitable statistical analyses, we could not reach firm conclusions regarding the effectiveness of low beta NFT itself. However, our analysis cautiously suggests that SMR NFT is more promising that low beta one.

Overall, our qualitative analysis on younger adults NFT protocols indicated that the amount of NFT, measured as total minutes of training, might be one of the success predictors and 120 min of NFT represented the minimum duration of training required to achieve significant electrophysiological and behavioral results. However, it is not the only factor involved. If so, the longest NFT protocols (i.e., Cannon et al., 2009; Kober et al., 2017) should have been successful, which was not the case.

Secondly, the three successful NFT protocols used single-band modulations; that is, they targeted a single frequency at a time. This might suggest that single band training protocols facilitate participants' learning, as suggested by Rogala and colleagues' (2016) review. Of course, future studies directly comparing single-band and multiband NFT modulation are highly recommended in order to gain more solid experimentally grounded insights than those based on this qualitative review of the few studies available.

Lastly, due to the multifaceted nature of executive functions, it is very likely that not just one frequency training is effective in producing executive function enhancements and our analysis confirmed this assumption, suggesting that each protocol could be generally associated with different executive functions. In addition, the nature of the behavioral task used to make pretraining versus posttraining comparisons might also play a role. More in detail, taking Enriquez-Geppert, Huster, Figge and colleagues (2014) and Wang and Hsieh (2013) as instances, they used similar protocols but obtained contrasting effects on executive function performance, as the former showed enhancements on proactive control tasks, whereas the latter improved reactive control tasks. Thus, it seems that employing different executive function tasks influenced the results, suggesting that the behavioral pretraining versus posttraining assessment should include a variety of executive function measures, ideally by also solving task impurity issues (e.g., Burgess, 1997; Miyake et al., 2000; Vallesi, 2020), in order to tap this multifaceted construct with more fine-grained precision.

Compared to younger adults' protocols, older adults' ones allowed us to reach even less clear conclusions. The main limitation was the low number of studies involving older adult participants, which did not permit to make generalizations. However, with this caveat in mind, based on the retrieved studies, we found that only theta NFT achieved positive outcomes (Figure 4). By contrast, we could not put forward any interpretation regarding alpha protocol results. In fact, although they seemed to be ineffective in the enhancement of older adults' executive functions, in Reis and collaborators' (2016) study, it was not possible to distinguish between the contribution of alpha and theta NFT, and in Gomez-Pilar and colleagues' (2016) study participants were required to downregulate this frequency. Similarly, we found no evidence of low beta efficacy and executive function improvements in older adults. Overall, what is clear is that more studies with older adults are needed to reach some conclusions.

Thus far, the discussion of the qualitative analysis results has dealt with our preliminary and tentative speculations and the most noticeable observation that emerged was a generally low quality of the NFT protocols, whose findings seemed to be sometimes based on inaccurate methodology and statistical analyses. Thus, since our most frequent criticism regarded the methodological drawbacks of the studies, the ROBINS – I (Sterne et al., 2016) and CRED–nf checklist (Ros et al., 2020) tools helped us to quantify study quality more precisely. With ROBINS – I, we evaluated whether the results of the analyzed NFT interventions could have been affected by biases. Figure 2 details the risk of bias for each domain, and, as can be seen, all studies were at low risk of bias in the two pre-intervention domains ("bias due to confounding" and "bias in selection of participants into the studies") and in the at-intervention domain ("bias in classification of interventions"). The low risk of bias in these three domains, which addresses possible issues before the start of the training, indicates that there was no confounding and that the intervention was well defined. The remaining four domains concern issues after the start of the intervention and appear to be the most critical. Most of the studies did not exhibit any risk of bias in the fourth domain, dealing with deviations from intended interventions. However, four studies made an exception and were at serious risk of bias in the fourth domain. More in details, the four studies were those in which the experimental group was compared to a passive control group, that is, control participants did not receive sham NFT but only a behavioral training or nothing at all in between pretraining and posttraining assessments. This procedure is at risk of producing the so-called "effect of assignment to intervention", as there might have been deviations due to participants' expectations, which were likely to depend on the group they belonged to. Greater similarity between studies was observed for the last three domains, but in the negative sense, as there was no study at low risk of bias. More in detail, all studies suffered from serious risk of bias due to missing data (domain 5), as the analysis was unlikely to have removed risk of bias from missing data. Moreover, Escolano and colleagues' (2011) study was at critical risk of bias in this domain, since the analysis of electrophysiological effects was performed excluding a priori nonresponder participants and no appropriate analysis addressed this issue. The scenario for the sixth domain, that is, bias in measurement of outcomes, is mixed. Studies at moderate risk of bias were generally those in which measurement of the outcome was appropriate and comparable across intervention groups and the outcome assessment was slightly influenced by the knowledge of the intervention received by participants. Serious risk of bias was introduced either if the method of outcome assessment was not comparable across groups (i.e., there was no sham control group) or if there was a systematic error in the assessment of the outcome (i.e., the interaction was not tested), whereas the risk of bias was critical if both the previous conditions were present. Lastly, the seventh domain addresses possible issues deriving from the selection of the reported results. In the studies at moderate risk of bias there was neither indication of selection of the reported analysis among multiple analyses nor selection of subgroups of participants, whereas in those at serious risk of bias there was high risk that analyses or subgroups of participants had been selected. Only one study was at critical risk of bias for this domain, suggesting that there was suspicion of selective reporting of results. The overall judgment suggests that all the included studies were at serious risk of bias and that three were at critical risk of bias, confirming our impression of methodological limitations in the analyzed studies. Moreover, the overall picture lends support to what we discussed previously. Indeed, it appears that studies for which we highlighted a higher number of drawbacks are also those at higher risk of bias, in terms of more domains at serious or critical risk. Lastly, in our view, the risk of bias assessment suggests that the reported effect sizes, which ranged from medium to very large, might be inflated as the studies were probably underpowered in terms of sample sizes.

The quality of the NFT protocols as assessed using the CRED–nf checklist also deserves some discussion (see Figure 3). It is interesting to note that the three "successful" studies were those that obtained a higher score in this checklist (80%–87%), probably indicating the good quality of their NFT protocols and, thus, confirming the reliability of their positive results. Other two "partially successful" studies were rated highly (Hsueh et al., 2016; Wei et al., 2017), and this might suggest that high-quality studies were not able to provide evidence for alpha NFT effectiveness. In line with the risk of bias assessment, the studies at higher risk of bias were almost the same as the ones that achieved the lowest percentages on the CRED–nf checklist (53%–60%). The studies in between might instead indicate the need of some enhancements at the methodological level to improve the reliability of their results.

In conclusion, the present systematic review did not identify robust evidence for NFT as an effective technique for enhancing healthy subjects' executive functions, as most of the analyzed studies did not achieve significant results at both the electrophysiological and behavioral levels. Nevertheless, the few studies pointing to its efficacy might reflect a promising starting point for future studies. Therefore, our results highlight that, to reliably verify whether NFT is effective at enhancing executive functions in healthy participants, further studies are definitely required to solve the encountered methodological issues and to provide a clearer causal relation.

4.1 | Suggestions for future studies

Our qualitative analysis, by underling the weaknesses of the retrieved studies, wishes to stress the importance of further well-controlled NFT studies specifically focusing on boosting executive functioning in the adult population. The paucity of rigorous studies can be noticed from the fact that our database search identified 62 eligible studies, but we could include only 20 of them in our qualitative synthesis which, in addition, were extremely heterogeneous and allowed limited comparisons. Hence, we will put forward some suggestions that might be useful for future studies.

Firstly, more controlled pre-NFT versus post-NFT designs would allow to identify electrophysiological and executive function changes induced by the training. Of primary concern is to use shared protocols with similar EEG measures. Moreover, concerning the pretraining versus posttraining behavioral assessment, it could be advisable to use tasks tapping on the same functions but differentiated in a sort of parallel forms in order to minimize learning effects in the posttraining assessment.

The second crucial issue regards the need of more rigorously chosen control groups in order to obtain more reliable training evidence. In our analysis, we found a large heterogeneity of control groups, which sometimes were passive control groups, not allowing reliable comparisons and control of confounding variables. Thus, the most desirable option would be to include a sham NFT group or to compare the NFT experimental group to another NFT group receiving feedback for an unrelated target-frequency.

Third, to provide unambiguous evidence for executive function enhancement, it is fundamental that the statistical analysis directly compares the experimental NFT group with a control group and reports the statistical significance of their difference. Comparing pretraining measures with posttraining ones in the experimental group per se does not allow to verify if it outperforms the control group, thus, if the effect is specifically produced by NFT.

Lastly, studies should be based on the same executive function operational definitions in order to define which target-frequencies are better to use to enhance specific executive functions. Moreover, based on the operational definition, studies should include a wider variety of executive function tasks in order to measure more in depth executive functions. In fact, more rigorous behavioral assessments could identify more precisely also subtle executive function enhancements.

ACKNOWLEDGMENTS

This work was supported by the "Department of excellence 2018–2022" initiative of the Italian Ministry of education (MIUR) awarded to the Department of Neuroscience–University of Padua. The authors thank Ettore Ambrosini for fruiful discussions during the preparation of this manuscript.

AUTHOR CONTRIBUTIONS

Giada Viviani: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Writing-original draft; Writing-review & editing. **Antonino Vallesi:** Conceptualization; Data curation; Formal analysis; Funding acquisition; Methodology; Project administration; Resources; Software; Supervision; Validation; Writing-original draft; Writing-review & editing.

ORCID

Antonino Vallesi https://orcid.org/0000-0002-4087-2845

REFERENCES

Angelakis, E., Stathopoulou, S., Frymiare, J. L., Green, D. L., Lubar, J. F., & Kounios, J. (2007). EEG neurofeedback: A brief overview and an example of peak alpha frequency training for cognitive enhancement in the elderly. *The Clinical Neuropsychologist*, *21*(1), 110–129. https://doi.org/10.1080/13854040600744839

Arns, M., Batail, J.-M., Bioulac, S., Congedo, M., Daudet, C., Drapier, D., Fovet, T., Jardri, R., Le-Van-Quyen, M., Lotte, F., Mehler, D., Micoulaud-Franchi, J.-A., Purper-Ouakil, D., & Vialatte, F. (2017). Neurofeedback: One of today's techniques in psychiatry? *L'encephale*, *43*(2), 135–145. https://doi.org/10.1016/j.encep.2016.11.003

Arns, M., de Ridder, S., Strehl, U., Breteler, M., & Coenen, A. (2009). Efficacy of neurofeedback treatment in ADHD: The effects on inattention, impulsivity and hyperactivity: A meta-analysis. *Clinical EEG and Neuroscience*, *40*(3), 180–189. https://doi.org/10.1177/155005940904000311

Barkley, R. A. (2012). *Executive functions: What they are, how they work and why they evolved*. The Guilford Press.

Berger, A. M., & Davelaar, E. J. (2018). Frontal alpha oscillations and attentional control: A virtual reality neurofeedback study. *Neuroscience*, *378*, 189–197. https://doi.org/10.1016/j.neuroscience.2017.06.007

Berger, H. (1930). Über das Elektrenkephalogramm des Menschen, 2nd report. *Journal für Psychologie und Neurologie*, *40*, 160–179.

Burgess, P. W. (1997). Theory and methodology in executive function research. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 81–116). Taylor & Francis.

Campos da Paz, V. K., Garcia, A., da Paz, C., Neto, A., & Tomaz, C. (2018). SMR neurofeedback training facilitates working memory performance in healthy older adults: A behavioral and EEG study. *Frontiers in Behavioral Neuroscience*, *12*, 321. https://doi.org/10.3389/fnbeh.2018.00321

Cannon, R., Congedo, M., Lubar, J., & Hutchens, T. (2009). Differentiating a network of executive attention: Loreta neurofeedback in anterior cingulate and dorsolateral prefrontal cortices. *International Journal of Neuroscience*, *119*(3), 404–441. https://doi.org/10.1080/00207450802480325

Cavanagh, J. F., Zambrano-Vazquez, L., & Allen, J. J. B. (2012). Theta lingua franca: A common mid-frontal substrate for action monitoring processes. *Psychophysiology*, *49*(2), 220–238. https://doi.org/10.1111/j.1469-8986.2011.01293.x

Coben, R., Linden, M., & Myers, T. E. (2010). Neurofeedback for autistic spectrum disorder: A review of the literature. *Applied Psychophysiology and Biofeedback*, *35*(1), 83–105. https://doi.org/10.1007/s10484-009-9117-y

Cohen, M. X., & Donner, T. H. (2013). Midfrontal conflict-related theta-band power reflects neural oscillations that predict behavior. *Journal of Neurophysiology*, *110*(12), 2752–2763. https://doi.org/10.1152/jn.00479.2013

Cooper, N. R., Croft, R. J., Dominey, S. J. J., Burgess, A. P., & Gruzelier, J. H. (2003). Paradox lost? Exploring the role of alpha oscillations during externally vs. internally directed attention and the implications for idling and inhibition hypotheses. *International Journal of Psychophysiology*, *47*(1), 65–74. https://doi.org/10.1016/s0167-8760(02)00107-1

Corydon Hammond, D., Bodenhamer-Davis, G., Gluck, G., Stokes, D., Hunt Harper, S., Trudeau, D., MacDonald, M., Lunt, J., & Kirk, L. (2011). Standards of practice for neurofeedback and neurotherapy: A position paper of the international society for neurofeedback & research. *Journal of Neurotherapy*, *15*(1), 54–64. https://doi.org/10.1080/10874208.2010.545760

de Zambotti, M., Bianchin, M., Magazzini, L., Gnesato, G., & Angrilli, A. (2012). The efficacy of EEG neurofeedback aimed at enhancing sensory-motor rhythm theta ratio in healthy subjects. *Experimental Brain Research*, *221*(1), 69–74. https://doi.org/10.1007/s00221-012-3148-y

Dempster, T., & Vernon, D. (2009). Identifying indices of learning for alpha neurofeedback training. *Applied Psychophysiology and Biofeedback*, *34*(4), 309–328. https://doi.org/10.1007/s10484-009-9112-3

Doppelmayr, M., & Weber, E. (2011). Effects of SMR and theta/beta neurofeedback on reaction times, spatial abilities, and creativity. *Journal of Neurotherapy*, *15*(2), 115–129. https://doi.org/10.1080/10874208.2011.570689

Egner, T., & Gruzelier, J. H. (2001). Learned self-regulation of EEG frequency components affects attention and event-related brain potentials in humans. *NeuroReport*, *12*(18), 4155–4159. https://doi.org/10.1097/00001756-200112210-00058

Egner, T., & Gruzelier, J. H. (2004). EEG Biofeedback of low beta band components: Frequency-specific effects on variables of attention and event-related brain potentials. *Clinical Neurophysiology*, *115*(1), 131–139. https://doi.org/10.1016/S1388-2457(03)00353-5

Egner, T., Zech, T. F., & Gruzelier, J. H. (2004). The effects of neurofeedback training on the spectral topography of the electroencephalogram. *Clinical Neurophysiology*, *115*(11), 2452–2460. https://doi.org/10.1016/j.clinph.2004.05.033

Engelbregt, H. J., Keeser, D., van Eijk, L., Suiker, E. M., Eichhorn, D., Karch, S., Deijen, J. B., & Pogarell, O. (2016). Short and long-term effects of sham-controlled prefrontal EEG-neurofeedback training in healthy subjects. *Clinical Neurophysiology*, *127*(4), 1931–1937. https://doi.org/10.1016/j.clinph.2016.01.004

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*, 420. https://doi.org/10.3389/fnbeh.2014.00420

Enriquez-Geppert, S., Huster, R. J., & Herrmann, C. S. (2013). Boosting brain functions: Improving executive functions with behavioral training, neurostimulation, and neurofeedback. *International Journal of Psychophysiology*, *88*(1), 1–16. https://doi.org/10.1016/j.ijpsycho.2013.02.001

Enriquez-Geppert, S., Huster, R. J., & Herrmann, C. S. (2017). EEG-neurofeedback as a tool to modulate cognition and behavior: A review tutorial. *Frontiers in Human Neuroscience*, *11*, 51. https://doi.org/10.3389/fnhum.2017.00051

Enriquez-Geppert, S., Huster, R. J., Scharfenort, R., Mokom, Z. N., Vosskuhl, J., Figge, C., Zimmermann, J., & Herrmann, C. S. (2013). The morphology of midcingulate cortex predicts frontal-midline theta neurofeedback success. *Frontiers in Human Neuroscience*, *7*, 453. https://doi.org/10.3389/fnhum.2013.00453

Enriquez-Geppert, S., Huster, R. J., Scharfenort, R., Mokom, Z. N., Zimmermann, J., & Herrmann, C. S. (2014). Modulation of frontal-midline theta by neurofeedback. *Biological Psychology*, *95*, 59–69. https://doi.org/10.1016/j.biopsycho.2013.02.019

Escolano, C., Aguilar, M., & Minguez, J. (2011). EEG-based upper alpha neurofeedback training improves working memory performance. *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, *2011*, 2327–2330. https://doi.org/10.1109/IEMBS.2011.6090651

Escolano, C., Navarro-Gil, M., Garcia-Campayo, J., & Minguez, J. (2014). The effects of a single session of upper alpha neurofeedback for cognitive enhancement: A sham-controlled study. *Applied Psychophysiology and Biofeedback*, *39*(3–4), 227–236. https://doi.org/10.1007/s10484-014-9262-9

Fingelkurts, A. A., & Fingelkurts, A. A. (2014). EEG oscillatory states: Universality, uniqueness and specificity across healthy-normal, altered and pathological brain conditions. *PLoS One*, *9*(2), e87507. https://doi.org/10.1371/journal.pone.0087507

Freunberger, R., Werkle-Bergner, M., Griesmayr, B., Lindenberger, U., & Klimesch, W. (2011). Brain oscillatory correlates of working memory constraints. *Brain Research*, *1375*, 93–102. https://doi.org/10.1016/j.brainres.2010.12.048

Gevins, A., Smith, M. E., McEvoy, L., & Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice. *Cerebral Cortex*, *7*(4), 374–385. https://doi.org/10.1093/cercor/7.4.374

Gomez-Pilar, J., Corralejo, R., Nicolas-Alonso, L. F., Álvarez, D., & Hornero, R. (2016). Neurofeedback training with a motor imagery-based BCI: Neurocognitive improvements and EEG changes in the elderly. *Medical & Biological Engineering & Computing*, *54*(11), 1655–1666. https://doi.org/10.1007/s11517-016-1454-4

Gonçalves, Ó. F., Carvalho, S., Mendes, A. J., Leite, J., & Boggio, P. S. (2018). Neuromodulating attention and mind-wandering processes with a single session real time EEG. *Applied Psychophysiology and Biofeedback*, *43*(2), 143–151. https://doi.org/10.1007/s10484-018-9394-4

Gordon, S., Todder, D., Deutsch, I., Garbi, D., Alkobi, O., Shriki, O., Shkedy-Rabani, A., Shahar, N., & Meiran, N. (2020). Effects of neurofeedback and working memory-combined training on executive functions in healthy young adults. *Psychological Research Psychologische Forschung*, *84*(6), 1586–1609. https://doi.org/10.1007/s00426-019-01170-w

Groppe, D. M., Bickel, S., Keller, C. J., Jain, S. K., Hwang, S. T., Harden, C., & Mehta, A. D. (2013). Dominant frequencies of resting human brain activity as measured by the electrocorticogram. *NeuroImage*, *79*, 223–233. https://doi.org/10.1016/j.neuroimage.2013.04.044

Grunwald, M., Weiss, T., Krause, W., Beyer, L., Rost, R., Gutberlet, I., & Gertz, H. J. (2001). Theta power in the EEG of humans during ongoing processing in a haptic object recognition task. *Brain Research. Cognitive Brain Research*, *11*(1), 33–37. https://doi.org/10.1016/s0926-6410(00)00061-6

Gruzelier, J. H. (2014). EEG-neurofeedback for optimising performance. I: A review of cognitive and affective outcome in healthy participants. *Neuroscience and Biobehavioral Reviews*, *44*, 124–141. https://doi.org/10.1016/j.neubiorev.2013.09.015

Gruzelier, J., Inoue, A., Smart, R., Steed, A., & Steffert, T. (2010). Acting performance and flow state enhanced with sensory-motor rhythm neurofeedback comparing ecologically valid immersive VR and training screen scenarios. *Neuroscience Letters*, *480*(2), 112–116. https://doi.org/10.1016/j.neulet.2010.06.019

Hanslmayr, S., Sauseng, P., Doppelmayr, M., Schabus, M., & Klimesch, W. (2005). Increasing individual upper alpha power by neurofeedback improves cognitive performance in human subjects. *Applied Psychophysiology and Biofeedback*, *30*(1), 1–10. https://doi.org/10.1007/s10484-005-2169-8

Hsueh, J., Chen, T., Chen, J., & Shaw, F. (2016). Neurofeedback training of EEG alpha rhythm enhances episodic and working memory. *Human Brain Mapping*, *37*(7), 2662–2675. https://doi.org/10.1002/hbm.23201

Jensen, O., & Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *The European Journal of Neuroscience*, *15*(8), 1395–1399. https://doi.org/10.1046/j.1460-9568.2002.01975.x

Jirayucharoensak, S., Israsena, P., Pan-Ngum, S., Hemrungrojn, S., & Maes, M. (2019). A game-based neurofeedback training system to enhance cognitive performance in healthy elderly subjects and in patients with amnestic mild cognitive impairment. *Clinical Interventions in Aging*, *14*, 347–360. https://doi.org/10.2147/CIA.S189047

Jurewicz, K., Paluch, K., Kublik, E., Rogala, J., Mikicin, M., & Wróbel, A. (2018). EEG-neurofeedback training of beta band (12–22 Hz) affects alpha and beta frequencies—A controlled study of a healthy population. *Neuropsychologia*, *108*, 13–24. https://doi.org/10.1016/j.neuropsychologia.2017.11.021

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews*, *29*(2–3), 169–195. https://doi.org/10.1016/s0165-0173(98)00056-3

Klimesch, W., Doppelmayr, M., Schwaiger, J., Auinger, P., & Winkler, T. (1999). "Paradoxical" alpha synchronization in a memory task. *Brain Research. Cognitive Brain Research*, *7*(4), 493–501. https://doi.org/10.1016/s0926-6410(98)00056-1

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition-timing hypothesis. *Brain Research Reviews*, *53*(1), 63–88. https://doi.org/10.1016/j.brainresrev.2006.06.003

Kober, S. E., Witte, M., Neuper, C., & Wood, G. (2017). Specific or nonspecific? Evaluation of band, baseline, and cognitive specificity of sensorimotor rhythm- and gamma-based neurofeedback. *International Journal of Psychophysiology*, *120*, 1–13. https://doi.org/10.1016/j.ijpsycho.2017.06.005

Lacroix, J. M. (1986). Mechanisms of biofeedback control. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (Vol. 4, pp. 137–162). Springer.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *PLoS Med*, *6*(7), e1000100. https://doi.org/10.1371/journal.pmed.1000100

Lofthouse, N., Arnold, L. E., Hersch, S., Hurt, E., & DeBeus, R. (2012). A review of neurofeedback treatment for pediatric ADHD. *Journal of Attention Disorders*, *16*(5), 351–372. https://doi.org/10.1177/1087054711427530

MacPherson, S. E., Della Sala, S., Cox, S. R., Girardi, A., & Iveson, M. H. (2015). *Handbook of frontal lobe assessment*. Oxford University Press.

MacPherson, S. E., Gillebert, C. R., Robinson, G. A., & Vallesi, A. (2019). Editorial: Intra- and inter-individual variability of executive functions: Determinant and modulating factors in healthy and pathological conditions. *Frontiers in Psychology*, *10*, 432. https://doi.org/10.3389/fpsyg.2019.00432

Mahjoory, K., Cesnaite, E., Hohlefeld, F. U., Villringer, A., & Nikulin, V. V. (2019). Power and temporal dynamics of alpha oscillations at rest differentiate cognitive performance involving sustained and phasic cognitive control. *NeuroImage*, *188*, 135–144. https://doi.org/10.1016/j.neuroimage.2018.12.001

Micoulaud-Franchi, J.-A., McGonigal, A., Lopez, R., Daudet, C., Kotwas, I., & Bartolomei, F. (2015). Electroencephalographic neurofeedback: Level of evidence in mental and brain disorders and suggestions for good clinical practice. *Clinical Neurophysiology*, *45*(6), 423–433. https://doi.org/10.1016/j.neucli.2015.10.077

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167

Mischel, W., Ayduk, O., Berman, M. G., Casey, B. J., Gotlib, I. H., Jonides, J., Kross, E., Teslovich, T., Wilson, N. L., Zayas, V., & Shoda, Y. (2011). 'Willpower' over the life span: Decomposing self-regulation. *Social Cognitive and Affective Neuroscience*, *6*(2), 252–256. https://doi.org/10.1093/scan/nsq081

Mitchell, D. J., McNaughton, N., Flanagan, D., & Kirk, I. J. (2008). Frontal-midline theta from the perspective of hippocampal "theta". *Progress in Neurobiology*, *86*(3), 156–185. https://doi.org/10.1016/j.pneurobio.2008.09.005

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Naas, A., Rodrigues, J., Knirsch, J.-P., & Sonderegger, A. (2019). Neurofeedback training with a low-priced EEG device leads to faster alpha enhancement but shows no effect on cognitive performance: A single-blind, sham-feedback study. *PLoS One*, *14*(9), e0211668. https://doi.org/10.1371/journal.pone.0211668

Nieuwenhuis, S., Forstmann, B. U., & Wagenmakers, E.-J. (2011). Erroneous analyses of interactions in neuroscience: A problem of significance. *Nature Neuroscience*, *14*(9), 1105–1107. https://doi.org/10.1038/nn.2886

Nigbur, R., Ivanova, G., & Stürmer, B. (2011). Theta power as a marker for cognitive interference. *Clinical Neurophysiology*, *122*(11), 2185–2194. https://doi.org/10.1016/j.clinph.2011.03.030

Nigro, S. E. (2019). The efficacy of neurofeedback for pediatric epilepsy. *Applied Psychophysiology and Biofeedback*, *44*(4), 285–290. https://doi.org/10.1007/s10484-019-09446-y

Omejc, N., Rojc, B., Battaglini, P. P., & Marusic, U. (2019). Review of the therapeutic neurofeedback method using electroencephalography: EEG neurofeedback. *Bosnian Journal of Basic Medical Sciences*, *19*(3), 213–220. https://doi.org/10.17305/bjbms.2018.3785

Pei, G., Wu, J., Chen, D., Guo, G., Liu, S., Hong, M., & Yan, T. (2018). Effects of an integrated neurofeedback system with dry electrodes: EEG acquisition and cognition assessment. *Sensors (Basel, Switzerland)*, *18*(10), 3396. https://doi.org/10.3390/s18103396

Reis, J., Portugal, A. M., Fernandes, L., Afonso, N., Pereira, M., Sousa, N., & Dias, N. S. (2016). An Alpha and theta intensive and short neurofeedback protocol for healthy aging working-memory training. *Frontiers in Aging Neuroscience*, *8*, 157. https://doi.org/10.3389/fnagi.2016.00157

Rogala, J., Jurewicz, K., Paluch, K., Kublik, E., Cetnarski, R., & Wróbel, A. (2016). The do's and don'ts of neurofeedback training: A review of the controlled studies using healthy adults. *Frontiers in Human Neuroscience*, *10*, 301. https://doi.org/10.3389/fnhum.2016.00301

Ros, T., Enriquez-Geppert, S., Zotev, V., Young, K. D., Wood, G., Whitfield-Gabrieli, S., Wan, F., Vuilleumier, P., Vialatte, F., Van De Ville, D., Todder, D., Surmeli, T., Sulzer, J. S., Strehl, U., Sterman, M. B., Steiner, N. J., Sorger, B., Soekadar, S. R., Sitaram, R., … Thibault, R. T. (2020). Consensus on the reporting and experimental design of clinical and cognitive-behavioural neurofeedback studies (CRED-nf checklist). *Brain*, *143*(6), 1674–1685. https://doi.org/10.1093/brain/awaa009

Ros, T., J. Baars, B., Lanius, R. A., & Vuilleumier, P. (2014). Tuning pathological brain oscillations with neurofeedback: A systems neuroscience framework. *Frontiers in Human Neuroscience*, *8*, 1008. https://doi.org/10.3389/fnhum.2014.01008

Sterman, M. B. (1996). Physiological origins and functional correlates of EEG rhythmic activities: Implications for self-regulation. *Biofeedback & Self Regulation*, *21*(1), 3–33. https://doi.org/10.1007/BF02214147

Sterne, J. A. C., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., Henry, D., Altman, D. G., Ansari, M. T., Boutron, I., Carpenter, J. R., Chan, A.-W., Churchill, R., Deeks, J. J., Hróbjartsson, A., Kirkham, J., Jüni, P., Loke, Y. K., Pigott, T. D., … Higgins, J. P. T. (2016). ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*, *355*, i4919. https://doi.org/10.1136/bmj.i4919

Tan, G., Thornby, J., Hammond, D. C., Strehl, U., Canady, B., Arnemann, K., & Kaiser, D. A. (2009). Meta-analysis of EEG biofeedback in treating epilepsy. *Clinical EEG and Neuroscience*, *40*(3), 173–179. https://doi.org/10.1177/155005940904000310

Thompson, M. M. D., & Thompson, L. (2003). *The neurofeedback book 2nd Edition: An introduction to basic concepts in applied psychophysiology*. AbeBooks.

Vallesi, A. (2020). The quest for hemispheric asymmetries supporting and predicting executive functioning. *Journal of Cognitive Neuroscience*, 1–19. https://doi.org/10.1162/jocn_a_01646

Vallesi, A., Tronelli, V., Lomi, F., & Pezzetta, R. (2021). Age differences in sustained attention tasks: A meta-analysis. *Psychonomic Bulletin & Review*. Advance Online Publication. https://doi.org/10.3758/s13423-021-01908-x

Vasquez, M., Gadea, M., Garijo, E., Aliño, M., & Salvador, A. (2015). Effects of assisted training with neurofeedback on EEG measures, executive function and mood in a healthy sample. *Anales de Psicologia*, *31*(1), 317–323. https://doi.org/10.6018/analesps.31.1.167241

Vernon, D. J. (2005). Can neurofeedback training enhance performance? An evaluation of the evidence with implications for future research. *Applied Psychophysiology and Biofeedback*, *30*(4), 347–364. https://doi.org/10.1007/s10484-005-8421-4

Vernon, D., Egner, T., Cooper, N., Compton, T., Neilands, C., Sheri, A., & Gruzelier, J. (2003). The effect of training distinct neurofeedback protocols on aspects of cognitive performance. *International Journal of Psychophysiology*, *47*(1), 75–85. https://doi.org/10.1016/s0167-8760(02)00091-0

Wang, J.-R., & Hsieh, S. (2013). Neurofeedback training improves attention and working memory performance. *Clinical Neurophysiology*, *124*(12), 2406–2420. https://doi.org/10.1016/j.clinph.2013.05.020

Wei, T.-Y., Chang, D.-W., Liu, Y.-D., Liu, C.-W., Young, C.-P., Liang, S.-F., & Shaw, F.-Z. (2017). Portable wireless neurofeedback system of EEG alpha rhythm enhances memory. *Biomedical Engineering Online*, *16*(1), 128. https://doi.org/10.1186/s12938-017-0418-8

West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin*, *120*(2), 272–292. https://doi.org/10.1037/0033-2909.120.2.272

Xiong, S., Cheng, C., Wu, X., Guo, X., Yao, L., & Zhang, J. (2014). Working memory training using EEG neurofeedback in normal young adults. *Bio-Medical Materials and Engineering*, *24*(6), 3637–3644. https://doi.org/10.3233/BME-141191

Yan, L., Wang, S., Yuan, Y., & Zhang, J. (2019). Effects of neurofeedback versus methylphenidate for the treatment of ADHD: Systematic review and meta-analysis of head-to-head trials. *Evidence-Based Mental Health*, *22*(3), 111–117. https://doi.org/10.1136/ebmental-2019-300088

Zanto, T. P., & Gazzaley, A. (2019). Aging of the frontal lobe. *Handbook of Clinical Neurology*, *163*, 369–389. https://doi.org/10.1016/B978-0-12-804281-6.00020-3

Zhang, G., Yao, L., Zhang, H., Long, Z., & Zhao, X. (2013). Improved working memory performance through self-regulation of dorsal lateral prefrontal cortex activation using real-time fMRI. *PLoS One*, *8*(8), e73735. https://doi.org/10.1371/journal.pone.0073735

Zoefel, B., Huster, R. J., & Herrmann, C. S. (2011). Neurofeedback training of the upper alpha frequency band in EEG improves cognitive performance. *NeuroImage*, *54*(2), 1427–1431. https://doi.org/10.1016/j.neuroimage.2010.08.078

How to cite this article: Viviani G, Vallesi A. EEG-neurofeedback and executive function enhancement in healthy adults: A systematic review. *Psychophysiology*. 2021;58:e13874. https://doi.org/10.1111/psyp.13874