Neuromodulation of brain activation associated with addiction: A review of real-time fMRI neurofeedback studies

Meghan E. Martz⁎, Tabatha Harta, Mary M. Heitzeg a, Scott J. Peltier b,c

a Addiction Center, Department of Psychiatry, University of Michigan, 4250 Plymouth Road, Ann Arbor, MI 48109, USA
b Functional MRI Laboratory, USA
c Department of Biomedical Engineering, Bonisteel Interdisciplinary Research Building, 2360 Bonisteel Blvd, Ann Arbor, MI 48109, USA

ARTICLE INFO

Keywords:
real-time fMRI
Neurofeedback
Addiction
Reward
Emotion

ABSTRACT

Real-time functional magnetic resonance imaging neurofeedback (rtfMRI-nf) has emerged in recent years as an imaging modality used to examine volitional control over targeted brain activity. rtfMRI-nf has also been applied clinically as a way to train individuals to self-regulate areas of the brain, or circuitry, involved in various disorders. One such application of rtfMRI-nf has been in the domain of addictive behaviors, including substance use. Given the pervasiveness of substance use and the challenges of existing treatments to sustain abstinence, rtfMRI-nf has been identified as a promising treatment tool. rtfMRI-nf has also been used in basic science research in order to test the ability to modulate brain function involved in addiction. This review focuses first on providing an overview of recent rtfMRI-nf studies in substance-using populations, specifically nicotine, alcohol, and cocaine users, aimed at reducing craving-related brain activation. Next, rtfMRI-nf studies targeting reward responsivity and emotion regulation in healthy samples are reviewed in order to examine the extent to which areas of the brain involved in addiction can be self-regulated using neurofeedback. We propose that future rtfMRI-nf studies could be strengthened by improvements to study design, sample selection, and more robust strategies in the development and assessment of rtfMRI-nf as a clinical treatment. Recommendations for ways to accomplish these improvements are provided. rtfMRI-nf holds much promise as an imaging modality that can directly target key brain regions involved in addiction, however additional studies are needed in order to establish rtfMRI-nf as an effective, and practical, treatment for addiction.

1. Introduction

Substance use impacts a multitude of lives across the world. In 2019, approximately 35 million people had a substance use disorder, and over half a million deaths were attributable to drug use (United Nations, 2019). In 2018, over 60% of Americans aged 12 and older reported past month use of substances such as alcohol, tobacco, or marijuana (Substance Abuse and Mental Health Services Administration, 2019). According to the United States Surgeon General's Report of Alcohol, Drugs, and Health (U.S. Department of Health and Human Services, 2016), alcohol misuse, illicit drug use, prescription drug use, and substance use disorders contributed to more than $400 billion costs associated with lost workplace productivity, health care expenses, and crime. A first step to reducing substance misuse, and its associated harms, is to better understand the underlying factors that contribute to addiction.

An abundance of evidence supports addiction as a brain disease (Leshner, 1997; Volkow et al., 2016). Consuming alcohol and drugs triggers an influx of dopamine in the brain's reward system (Volkow et al., 2017). The euphoric experience associated with this influx is reinforcing, which contributes to repetitive and compulsive substance use (Roob & Volkow, 2010). Advancements in neuroimaging techniques have allowed researchers to better understand brain systems involved in addiction, including those involved in positive and negative re-inforcement, decision making, and cognitive control (Ekhtiari et al., 2016). This work has also been used to develop neuroscience-informed substance use treatments (Chung et al., 2016).

One such neuroimaging advancement is real-time functional magnetic resonance imaging neurofeedback (rtfMRI-nf). The broad goal of neurofeedback is to train individuals how to self-regulate brain activity by providing real-time performance feedback. Traditional fMRI studies, which use non-invasive imaging to indirectly measure neuronal activity using the blood-oxygen-level-dependent (BOLD) signal in the brain, are typically used to acquire task-based or resting-state data that is then...
analyzed offline using region of interest (ROI) or whole-brain analyses. Although rtfMRI-nf studies also use the hemodynamic response from BOLD signal as a metric of brain function, they are unique from traditional fMRI studies by first training participants to control brain activation and then relaying brain activation from a particular brain region or network back to the participant through a dynamic feedback mechanism. Participants use this feedback to modify their behavior and resultant brain activation in order to achieve a desired outcome (e.g., decreasing a thermometer linked to craving-related neural activity).

Thus, neurofeedback is a form of operant conditioning (Skinner, 1938). An advantage of rtfMRI-nf is that regulation of the hemodynamic response occurs much more quickly than regulation of the electrical signals measured by EEG (Thibault et al., 2018). Furthermore, rtfMRI-nf can be used to target a much wider range of brain regions, such as subcortical reward regions central to neural processes involved in addiction, compared to other neurofeedback modalities which only capture surface cortical areas (see Stoeckel et al., 2014; Thibault et al., 2018 for in-depth reviews on the history and progress of neurofeedback research).

Due to the closed-loop design of rtfMRI-nf, it is vital that the brain signal used for neurofeedback is both accurate and reliable (Stoeckel et al., 2014). Target regions or circuitry used in rtfMRI-nf are usually determined \textit{a priori} by using either 1) atlas-based coordinates for ROIs based on prior literature demonstrating associations between such regions and phenomena of interest; or 2) a functional localizer to identify regions or circuitry involved in neurocognition related to the focus of the study. Functional localizers include existing fMRI tasks known to elicit neural activation in certain areas of the brain, such as the monetary incentive delay task activating the ventral striatum (VS), or determining which areas of the brain are activated while participants engage in certain cognitive behaviors, such as resisting the urge to smoke cigarettes. Studies using rtfMRI-nf often examine if neurofeedback improves the ability to modulate brain activation by comparing this ability during trials with and without neurofeedback. Studies typically define the former as “training runs” and the latter as “transfer runs”. Effects that extend beyond the neurofeedback training run indicate learned behavior. This learned behavior has important implications for clinical rtfMRI-nf treatments beyond the laboratory setting, such as addiction interventions (Fig. 1).

Indeed, rtfMRI-nf has been identified as a promising method for the treatment of addiction, both as a primary treatment approach and in conjunction with other behavioral interventions (Stoeckel et al., 2014). An important goal of behavioral treatments, such as cognitive behavioral therapy (CBT), is to help individuals better understand the interrelatedness of thoughts, behaviors, and emotions. By changing negative thoughts, emotions can improve, and problem behaviors can decrease. Aspects of neural function relevant to addiction, including reward responsivity, emotion regulation, and self-control, are directly related to the core features of CBT (Carroll & Kiluk, 2017). rtfMRI-nf may help to supplement such psychotherapies by training individuals to modulate brain function directly related to these core features. rtfMRI-nf may also be useful the reduction of cravings in individuals with substance use addiction. Although still debated in the literature, drug cravings are typically viewed as an unconscious process (Tiffany & Wray, 2012). Clinical rtfMRI-nf studies focused on reducing drug cravings often use drug cues to localize areas of craving-related brain activation and then instruct participants to reduce cravings by providing neurofeedback from those regions. In this way, rtfMRI-nf allows for individuals to exert control over unconscious brain function directly related to cognitions that sustain addiction. A subsequent goal of this application of rtfMRI-nf is to extend strategies to reduce drug cravings into real-life contexts outside of the scanner (Ekhtiari et al., 2016).

In this selective literature review, we provide an overview of rtfMRI-nf studies relevant to addiction published within the past 10 years. First, we review clinical rtfMRI-nf studies of substance use published within the past 10 years. These studies are organized by type of substance use (e.g., nicotine, alcohol, cocaine) targeted in these studies. Second, we review basic science rtfMRI-nf studies that used this imaging modality in order to better understand neural processes involved in addiction, specifically reward responsivity and emotion regulation. We then conclude with a discussion of possible future directions for the use of rtfMRI-nf as a clinical intervention and research tool relevant to substance use. For each clinical rtfMRI-nf study reviewed, we used “real-time fMRI”, “neurofeedback”, and either “alcohol”, “nicotine”, or “cocaine” as key search terms and required that at least one participant group in each study be comprised of heavy substance users or individuals in treatment for a substance use disorder. For basic science rtfMRI-nf studies selected for our review, we used “real-time fMRI”, “neurofeedback”, and either “reward” or “emotion regulation” as key search terms and required that at least one participant group consist of healthy participants. Fig. 2 provides a visualization of target brain regions used in the rtfMRI-nf studies we reviewed.

![Fig. 1. Ideal rtfMRI-nf study design for substance-using participants.](image-url)
2. Real-time fMRI neurofeedback research in substance-using samples

The rewarding attributes of drugs and alcohol tend to weaken the ability of individuals to exert self-control over substance use, often resulting in a cycle of craving and drug seeking behavior (Volkow & Morales, 2015). The course of addiction has been described in three interacting stages: 1) preoccupation/anticipation; 2) binge/intoxication; and 3) withdrawal/negative affect (Koob, 2009; Koob & Volkow, 2010). The first stage is characterized by decision-making processes and self-control over impulsive responding and primarily involves the prefrontal cortex (PFC). The second stage is the actual use experience, when substances cause an influx of dopamine in the brain’s reward circuitry. Dopamine neurons in the ventral tegmental area (VTA) project to the VS, which includes the nucleus accumbens (NAcc). The NAcc is sensitive to rewarding and novel stimuli and has reciprocal links to decision-making processes in the PFC. The third stage is characterized by negative affect associated with withdrawal. During this stage, negative affect involves functioning of the limbic system—including the amygdala, anterior cingulate cortex (ACC), and insula—and perpetuates drug use in order to restore the positive, albeit short-lived, rewarding effects of substances. Their cyclical interaction makes these three stages potential targets for clinical interventions. Clinically focused rtfMRI-nf studies typically use this imaging modality to train participants in behavioral self-regulation in order to reduce substance craving and, ultimately, the extent of participants’ substance use.

2.1. Nicotine

Tobacco use is the leading cause of preventable morbidity and mortality in the United States (U.S. Department of Health and Human Services, 2014). Nicotine, the ingredient in products such as cigarettes, electronic vaping products, and smokeless tobacco is highly addictive. After consumption, nicotine stimulates the release of epinephrine and activates reward circuitry in the brain, thereby increasing dopamine (National Institute on Drug Abuse, 2020). According to the 2020 Surgeon General’s report on smoking cessation (U.S. Department of Health and Human Services, 2020), over two-thirds of smokers have intentions to quit; however, it takes multiple attempts to sustain abstinence due to intense nicotine cravings often impeding intentions to quit. Therefore,
rtfMRI-nf studies using nicotine-dependent participants have examined the use of neurofeedback to reduce craving-related brain activation. Each of the nicotine-related rtfMRI-nf studies we reviewed primarily targeted the ACC (Canterberry et al., 2013), with the majority also examining activation in prefrontal regions involved in decision-making and goal-directed behaviors (Hanlon et al., 2013; Hartwell et al., 2016; Kim et al., 2015; Li et al., 2013). As a central hub between prefrontal control systems and reward systems, the ACC is involved in regulation of cognitive and emotional processing (Allman et al., 2001; Bush et al., 2000). Thus, ACC activation is believed to be central to craving, whereas PFC function is associated with resistance to craving.

Canterberry et al. (2013), Hanlon et al. (2013), Hartwell et al. (2016), and Li et al. (2013) each used similar rtfMRI-nf paradigms to examine the extent to which nicotine-dependent adults could reduce craving-related brain activation in response to smoking cues. Participants in each of these studies were nicotine-dependent adults residing in the southeastern United States. Participants were non-treatment seeking, community-recruited adults in Canterberry et al. (2013) and Hartwell et al. (2016) and treatment-seeking adults in Hanlon et al. (2013) and Li et al. (2013). As reported in Hanlon et al. (2013) and Li et al. (2013), participants first completed two ROI isolation runs in which participants were instructed to “allow yourself to crave” and “resist the urge to smoke”, respectively, while viewing smoking-related images. First, craving-related activation was localized in the ventral ACC (crave ROI), and activation associated with resistance to smoking was localized in the dorsomedial PFC (resist ROI). Then during neurofeedback runs, feedback from the crave ROI and resist ROI were presented to participants via veridical thermometers. Participants were instructed to decrease the crave thermometer and increase the resist thermometer.

Hanlon et al. (2013) used this study paradigm in three separate rtfMRI-nf visits, which were approximately 7 to 10 days apart. Findings showed that participants were able to successfully lower ventral ACC activation during crave trials when receiving neurofeedback compared to baseline. Self-reported craving ratings correlated with ACC activation; as participants successfully lowered ACC activation with rtfMRI-nf they also reported less subjective cravings. However, participants were not able to significantly modulate brain activity in the dorsomedial PFC during resist trials, even when receiving neurofeedback. Using the same rtfMRI-nf procedure as Hanlon et al. (2013), but during a single study visit, Li et al. (2013) reported that participants were able to significantly decrease craving-related ACC activation during neurofeedback. Like Hanlon et al. (2013), participants from the Li et al. (2013) study were unable to significantly increase their prefrontal activation while trying to resist the urge to smoke. Li et al. (2013) stated that participants showed a reduction in self-reported cravings. These studies demonstrate the potential validity of using rtfMRI-nf, specifically to down-regulate ACC activation, in nicotine addiction treatment.

Based on findings of a more robust effect of neurofeedback to craving-related ACC activation compared to PFC activation, Canterberry et al. (2013) targeted only the ACC with the aim of reducing nicotine craving during rtfMRI-nf sessions. Participants who reported less severe nicotine dependence were better able to use neurofeedback to decrease their ACC activation over the course of three visits. Importantly, decreases in ACC activation were associated with reductions in self-reported cravings after just the first study visit. Taken together, findings reported by Canterberry et al. (2013) suggest that the effectiveness of rtfMRI-nf in craving reduction may be dependent upon the severity of nicotine dependence.

Similar to Canterberry et al. (2013), Hartwell et al. (2016) did not use separate “crave” and “resist” runs in their rtfMRI-nf paradigm, instead instructing participants only to reduce craving-related brain activation. Hartwell et al. (2016) used individualized ROIs within the ACC, including the medial PFC and orbitofrontal cortex, based on participants’ craving-related brain activation that visit. As a randomized, controlled study, this study included non-treatment seeking, nicotine-dependent adults who received neurofeedback and a matched control group of smokers who did not. Results indicated that the neurofeedback group had a significant mean reduction in craving-related brain activation compared to the non-neurofeedback control group. The control group did, however, show a downward trend in their activation. Hartwell et al. (2016) also compared the neurofeedback and non-neurofeedback groups on post-scan subjective craving ratings. The neurofeedback group had a significant reduction in the urge to smoke in order to experience pleasurable outcomes, but no significant group differences were found for the urge to smoke in order to relieve negative affect. Results from this study demonstrate that rtfMRI-nf appears to reduce both craving-related brain activation, as well as subjective experiences of craving related specifically to positive affect, at least in the relative short-term. Additional work is needed to determine the long-term extent of these reductions.

In a sample of Korean, nicotine-dependent males, Kim et al. (2015) used a connectivity approach in their neurofeedback paradigm in order to more accurately target the multiple brain regions involved in nicotine craving. Two sets of ROIs (ROI set 1: anterior ROIs comprised of the bilateral ACC, medial PFC, and orbitofrontal cortex; ROI set 2: posterior ROIs comprised of the bilateral posterior cingulate cortex (PCC) and precuneus) were collected at the beginning of the scanning session based on their likely involvement in craving-related brain activation. Participants were randomized into two groups based on type of neurofeedback signal extraction: 1) rtfMRI-nf using brain activity from ROI set 1 and 2) functional connectivity (FC)-added rtfMRI-nf by calculating Pearson’s correlations between the average BOLD time series for the two sets of ROIs. Across two separate rtfMRI-nf visits, participants were instructed to resist the urge to smoke when presented with the smoking video clips. The rtfMRI-nf group received neurofeedback based on BOLD activation from ROIs1 and the FC-added rtfMRI-nf group received neurofeedback based on the combination of BOLD activation in ROIs1 and functional connectivity between ROIs1 and ROIs2. Findings indicated that participants in both conditions significantly increased overall neural activity in ROIs1 and ROIs2 from the first visit to the second, with the FC-added neurofeedback group showing an overall greater level of activation. Furthermore, only the FC-added rtfMRI-nf group had a significant association between subjective craving score and mean neuronal activity, although this was limited to anterior ROIs. Of note, this is the first study to use neurofeedback based on connectivity in a sample of substance using participants. The results illustrate the multiple brain regions involved in nicotine craving and demonstrate the potential value of incorporating functional connectivity with rtfMRI-nf in addiction treatment. Because the sample was relatively homogenous, it would be useful for future research to examine the validity and replicability of their study design in more diverse populations.

2.2. Alcohol

To our knowledge, two published rtfMRI-nf studies used alcohol-using samples. In exploratory pilot studies, Karch et al. (2015) and Kirsch et al. (2016) tested the feasibility of neurofeedback training to modulate brain activation related to alcohol craving. Both of these studies were conducted using German participants, with treatment-seeking patients with alcohol use disorder (AUD) compared with healthy controls (Karch et al., 2015), and heavy-drinking college students (no control group; Kirsch et al., 2016). These studies also used different study designs to test the utility of rtfMRI-nf in reducing craving-related brain activation to alcohol cues. During a single study visit, Karch et al. (2015) examined if rtfMRI-nf could be used to modulate regional activation related to alcohol craving and assessed if rtfMRI-nf would contribute to changes in functional connectivity in the ACC and other regions within default mode network (DMN). The study design included a resting state scan before and after rtfMRI-nf sessions in order to measure pre- and post-scan functional
connectivity. Both the AUD patient group and healthy control group were divided into a subset that received true neurofeedback and another subset that received sham neurofeedback. Based on activation during an alcohol-cue exposure paradigm completed just prior to the rtfMRI-nf paradigm, the AUD patient true neurofeedback group viewed a thermometer linked to individualized ROI activation (ACC, dorsolateral PFC, or insula), and the healthy control true neurofeedback group’s thermometer was linked to PFC activation. Both sham AUD patient and sham healthy control groups’ thermometers were linked to brain activation in the cuneus that was unrelated to alcohol craving. Patients with AUD who received true neurofeedback showed significantly lower ROI activation while viewing alcohol cues during neurofeedback training. However, there were considerable individual differences in this ability. The healthy control sample, as well as both sham neurofeedback groups, did not show significant changes in BOLD activation during neurofeedback training. Furthermore, areas within the DMN, specifically medial frontal areas and the dorsolateral PFC, had increased connectivity after rtfMRI-nf in the AUD patient true neurofeedback group, as well as greater connectivity between frontal and subcortical regions. Supported by subjective craving ratings pre- and post-training, patients in the AUD patient true neurofeedback group reported less alcohol craving following rtfMRI-nf. Although preliminary, these results provide compelling evidence that rtfMRI-nf may contribute to reduced alcohol craving-related brain activation in individual ROIs and brain circuitry after only a single session.

Given the central role of the brain’s reward system in relation to addiction vulnerability (Volkow et al., 2017), Kirsch et al. (2016) used a monetary incentive delay task to target activation in the VS for neuromodulation. The full sample of heavy alcohol-using participants were separated into three groups: 1) an experimental group that received neurofeedback from the VS; 2) a control group that received “yoke” neurofeedback from another participant; and 3) a control group that viewed alcohol images with no feedback provided. During a single scan session, participants completed neurofeedback blocks and a non-neurofeedback session, in which they were instructed to downregulate VS activity. Findings indicated that participants in the true neurofeedback condition had significantly greater downregulation of VS activity compared to the other groups. As indicated by a questionnaire pre- and post-scanning, the true neurofeedback group did not report increased craving to alcohol after the scanning session. The true neurofeedback group also displayed correlations between PFC and VS activation, whereas the other groups did not. The authors interpret their findings in light of effective strategies often used in addiction treatment, such as efforts to reduce reward saliency and strengthen self-regulation over reward-driven impulses. Post-hoc analyses showed that the most frequently reported down-regulation strategies were imagining negative consequences of alcohol use, distraction by using negative thoughts, distraction by using positive or neutral thoughts, and using mindfulness techniques to reduce alcohol related thoughts. Asking participants to report strategies used during neurofeedback, as done by Kirsch et al. (2016), may inform the context of addiction-related treatment curriculum, particularly interventions such as dialectical behavioral therapy that include mindfulness and emotion regulation components.

2.3. Cocaine

Kirschner et al. (2018) investigated the ability of cocaine users to modulate brain activation in the dopaminergic mesolimbic reward system, specifically the VTA and substantia nigra (SN), using rtfMRI-nf. In a sample of cocaine users and a sample of healthy matched controls, the study design consisted of an anatomical localizer scan to identify the VTA/SN and then neurofeedback runs using combined activation from the VTA/SN. Using a paradigm adapted from Sulzer et al. (2013b; described later in this review), participants completed four runs; a pre-training non-neurofeedback run, two neurofeedback runs, and a post-training non-neurofeedback transfer run. Participants were asked to use mental imagery to increase VTA/SN activity during “happy time” trials and decrease VTA/SN activation during rest trials for each run. There were no group differences in the ability to modulate VTA/SN activation during the pre-training and neurofeedback runs between cocaine users and healthy controls. However, when cocaine users were categorized by the presence or absence of severe obsessive–compulsive drug use, the former subgroup showed deficits in VTA/SN modulation compared to healthy controls. There were no group differences in transfer effects, although the authors note that longitudinal research designs may be necessary to investigate such effects. Findings from this study demonstrate that cocaine users are able to exert volitional control over VTA/SN activation using rtfMRI-nf, which could translate into clinical treatment for patients with cocaine use disorder. This study also adds to work conducted by Canterberry et al. (2013) that demonstrates an association between severity of drug use and the ability to exert volitional control over targeted brain activity using neurofeedback.

3. Clinically-relevant basic science research

Although the aforementioned studies have identified rtfMRI-nf as a promising clinical tool to modulate brain function associated with addiction, these studies have not fully established rtfMRI-nf as an efficacious therapy. Long-term impacts on the treatment of addiction using rtfMRI-nf remain unknown. Furthermore, clinical studies have predominately only targeted one type of drug use, limiting their generalizability to other substances. The majority of these studies are framed as exploratory pilot studies that used small, relatively homogenous samples. Although the rtfMRI-nf studies reviewed below did not use heavy substance using participants, the brain systems targeted in these basic science studies correspond directly to two key neural correlates of addiction: 1) reward responsivity (Greer et al., 2014; MacInnes et al., 2016; Sulzer et al., 2013b); and 2) emotion regulation (Caria et al., 2010; Cohen Kadosh et al., 2016; Hamilton et al., 2011; Koush et al., 2017; Lawrence et al., 2014; Zotev et al., 2011).

3.1. Neuromodulation of reward-related brain activation

The VTA is located within the midbrain and contains a greater number of dopaminergic neurons than anywhere else in the brain. As such, the VTA plays a central role in reward motivated behavior and learning (Volkow et al., 2017), MacInnes et al. (2016) examined the feasibility of VTA modulation through rtfMRI-nf and effects of neurofeedback training on functional connectivity, which was measured by z-score changes in Pearson’s correlations. Participants were randomized into four groups: 1) a rtfMRI-nf group with neurofeedback linked to the VTA (VTA neurofeedback); and three other groups: 2) one given a dynamic cue to maintain attention (visual control), 3) one group that received neurofeedback associated with random noise (false neurofeedback), and 4) another group that received neurofeedback from the NaCC (NaCC neurofeedback). The VTA neurofeedback group showed greater functional connectivity with mesolimbic networks, whereas the other groups did not show significant changes in connectivity. Because they included a post-test, a transfer run after neurofeedback training, MacInnes et al. (2016) were able to demonstrate that self-regulation of VTA activation extended past the training session during which no feedback was provided. Although this study did not test long-term transfer effects, findings indicate the potential of neurofeedback to produce sustained effects of reward-related activation in the VTA.

Due to their functional similarities and dense concentration of dopaminergic neurons, Sulzer et al. (2013b) used the combined VTA/SN complex in their study examining the ability to up-regulate VTA/SN activation through reward-related imagery during rtfMRI-nf. Participants in the true neurofeedback condition received feedback linked to BOLD signal in VTA/SN that rose vertically in response to reward-related cognitions. The control group received inverted neurofeedback; in this condition they were provided with the same instructions but...
received feedback that decreased in relation to increased VTA/SN activation. Results showed that both groups were able to up-regulate VTA/SN activation, but only the true neurofeedback group was able to significantly increase VTA/SN activation beyond the baseline level. This group also had increased functional connectivity within the ni-
grostriatal pathway compared to the control group. As with findings from MacInnes et al. (2016), Sulzer et al. (2013b)'s results suggest that brain regulation in the VTA is malleable, and thus, a potential target for clinical interventions aimed at reward-related self-regulation. Findings from MacInnes et al. (2016) and Sulzer et al. (2013a,b) also demonstrate a reward-based learning component of rtfMRI-nf. Indeed, Kirschner et al. (2018), which adapted the paradigm developed by Sulzer et al. (2013b) in their rtfMRI-nf study in cocaine users, found that neurofeedback enhanced learning to modulate VTA/SN activation. Such learning has implications for using rtfMRI-nf as a clinical tool to implement behavior change in various contexts beyond the laboratory, such as using certain thoughts or imagery to reduce reward-related responding when individuals engage in substance use-related decision-making. This use of cognition relates to therapeutic approaches aimed at linking patterns of thoughts and behaviors (e.g., CBT).

Neuronal projections from the VTA extend primarily into the NAcc, PFC, hippocampus, and amygdala. The NAcc demonstrates robust activation in response to the anticipation of rewards. As such, NAcc activation is particularly relevant to reward-motivated, drug seeking behavior. Greer et al. (2014) is the first study to test if individuals could up- and down-regulate NAcc activation through rtfMRI-nf. Instead of using control groups, this study compared differences in up- and down-regulation of NAcc activation between neurofeedback versus non-neurofeedback runs within subjects. They also used psychophysiological interaction analyses (PPI) to examine connectivity between the NAcc and medial PFC to test if the strength of functional connectivity between these anatomically connected structures increased during neurofeedback. They found that participants were able to both up- and down-regulate NAcc activation just by using cognitive strategies and imagery, but participants were most successful during increase trials. There was a significant correlation between medial PFC and NAcc activation during neurofeedback, but not during non-feedback trials, indicating that the medial PFC may be recruited to facilitate reward-based learning. As with the learning effects discussed in MacInnes et al. (2016) and Sulzer et al. (2013b), participants were able to sustain volitional control of brain activity in non-neurofeedback transfer runs. This work has promising implications for the ability of individuals to self-regulate reward activation in order to have greater control over substance use.

3.2. Neuromodulation of affective brain activation

Negative affect associated with amygdala activation characterizes a component of addiction in which individuals use substances in order to diminish symptoms of withdrawal (Koob, 2009; Koob & Volkow, 2010). In line with the internalizing pathway to substance use, individuals experiencing anxiety and depression may use substances as a way to cope with negative affect (Hardee et al., 2018; Hussong et al., 2011). Thus, using rtfMRI-nf to investigate if individuals can exert volitional control over brain regions involved in emotion regulation has direct clinical implications relevant to the treatment of substance use disorders. Because the majority of existing rtfMRI-nf studies focus on emotion regulation (Fede et al., 2020), our review covers studies most relevant to addiction.

As described earlier in this review, studies that included substance-using samples have often targeted the ACC due to its centrality to craving-related brain activation (e.g., Canterberry et al., 2013; Hanlon et al. 2013; Li et al. 2013). This work is based, to a certain extent, on research conducted by Hamilton et al. (2011) that demonstrated the feasibility of down-regulating ACC activation through rtfMRI-nf and transfer effects that maintained volitional control of the ACC without the aid of neurofeedback. Specifically, Hamilton et al. (2011) targeted the subgenual ACC, which is involved in the generation of affective states. Participants were instructed to down-regulate subgenual ACC activation using any strategies of their choosing to increase their mood. A true neurofeedback group received feedback from the subgenual ACC, whereas a control group received yoked neurofeedback that was linked to brain activity from the true neurofeedback group. Results showed that true neurofeedback contributed to a significant reduction in subgenual ACC activation, and participants in the control group were not able to down-regulate their subgenual ACC. The true neurofeedback group showed decreased functional connectivity, as measured by PPI analyses, between the subgenual ACC and PCC during neurofeedback versus non-neurofeedback sessions, while the control group showed an increase between subgenual ACC-PCC functional connectivity. The PCC is a central hub within the DMN and has been found to play an important role in attention regulation (Hampson et al., 2006) and internally directed cognitions (Leech & Sharp, 2014). Thus, individuals who display stronger recruitment of the PCC while trying to regulate the subgenual ACC may be better able to neuromodulate that area of the brain to better regulate affective responding. These regions may be useful targets in future rtfMRI-nf studies in addicted populations who report deficits in emotion regulation associated with their substance use.

In addition to the ACC, the amygdala is a region central to emotion processing and negative affect involved in addiction (Koob, 2009). Zotev et al. (2011) tested if rtfMRI-nf could be used to modulate activation in the amygdala by instructing participants to increase BOLD signal in the left amygdala by using positive autobiographical memories. One group received true amygdala neurofeedback and a control group received sham feedback from the intraparietal sulcus. Results indicated that the true neurofeedback group was able to increase brain activation in the amygdala both during neurofeedback and non-neurofeedback conditions. The latter finding suggests a transfer effect in which participants were able to maintain and extend strategies to increase amygdala activation without directly receiving feedback stimuli. Furthermore, there were significant network correlations within the fronto-temporo-limbic network and increased General Linear Model-based functional connectivity between the amygdala and right medial frontal polar cortex, bilateral dorsomedial PFC, left ACC, and bilateral superior frontal gyrus during both neurofeedback and transfer sessions of the rtfMRI-nf task. Because the true neurofeedback group demonstrated a greater extent of amygdala modulation compared to the control group, it is likely that group differences are attributable to neurofeedback-based learning effects rather than practice effects.

Also targeting amygdala modulation through rtfMRI-nf, Koush et al., 2017 developed a novel, dynamic causal modeling (DCM)-based rtfMRI-nf paradigm. During three sessions across three separate visits, the primary aim of this study was to train participants to use DCM-rtfMRI-nf to increase top-down connectivity from the dorsomedial PFC to the amygdala in response to positive social images. To test learning effects, participants were asked to self-regulate emotional responding in the absence of neurofeedback. Findings indicated that the DCM-rtfMRI-nf paradigm did indeed strengthen emotion regulation. Compared to the sham neurofeedback group, the group receiving true neurofeedback had increased top-down, and decreased bottom-up, connectivity between the dorsomedial PFC and amygdala. Both Koush et al. (2017), Zotev et al. (2011) note that determining individuals’ abilities to modulate amygdala activation, specifically down-regulating this activity, and the extent to which this learning translates to behavioral changes as important areas of future clinical research. This is relevant to clinical efforts to strengthen emotion-related self-control in the context to substance use behavior, particularly in adolescent populations who have weaker cognitive control over socioemotional responding compared to other age groups (Shulman et al., 2016; Steinberg, 2010).

Similar to Caria et al. (2010), Koush et al. (2017) Zotev et al. (2011)
## Table 1
Summary of reviewed rTfMRI-nf studies relevant to addiction.

| Study | Target ROI(s)/Circuitry | Study Design | Participants | Control Condition/Group | Key Findings | Follow-Up |
|-------|-------------------------|--------------|--------------|-------------------------|--------------|-----------|
| **Nicotine** | | | | | | |
| Canterberry et al. (2013) | ACC | Intermittent nf presented via veridical thermometer; 3 nf runs, decrease craving trials using cognitive strategies during smoking cues | N = 9 nicotine-dependent smokers, 18-60 years old, 11% female | Neutral cue condition | Decreased ACC activation and craving over 3 visits; less severe nicotine dependence associated with better neuromodulation | Yes, 3 visits |
| Hanlon et al. (2013) | Ventral ACC and dorsal medial PFC | Intermittent nf presented via 2 veridical thermometer, one linked to craving-related ROI activation and the other resistance to craving-related ROI activation; 3 nf runs, decrease craving and increase resist trials using cognitive strategies during smoking cues | N = 15 treatment-seeking nicotine-dependent smokers, 21-45 years old, sex not reported | Neutral cue condition | Better able to decrease craving-related ROI activation; correlation between ventral ACC activation and self-reported craving across visits | Yes, 3 visits |
| Hartwell et al. (2016) | ACC and medial PFC | Intermittent nf presented via veridical thermometer; 3 nf runs, decrease craving trials using cognitive strategies during smoking cues | N = 44 non-treatment seeking nicotine-dependent smokers, 18-60 years old; nf group n = 21, M age = 34.1, 38% female, non-nf group n = 23, M age = 36.2, 35% female | No-nf control group; neutral cue condition | Nf group decreased craving-related ROI activation and physiological and subjective craving across visits | Yes, 3 visits |
| Kim et al. (2015) | ACC, PFC, functional connectivity to PCC and precuneus | Continuous nf via opacity of smoking image cues; 6 nf runs, decrease craving trials using cognitive strategies; pre- and post-nf runs, fc | N = 10 nicotine-dependent smokers, ages 21-60 years old, M age = 28.7, 67% female | Neutral cue condition | Better able to decrease craving-related ACC activation; unable to increase resistance-related medial PFC activation | None |
| Li et al. (2013) | ACC and medial PFC | Intermittent nf via veridical thermometer; 2 nf training runs, decrease craving trials and increase resistance to craving using cognitive strategies during smoking cues | N = 14 non-treatment seeking nicotine-dependent smokers, all male; nf group n = 7, M age = 26.0, fc-added neurofeedback group n = 7, M age = 26.0 | Fc-added nf group better able to increase brain activity and had greater connectivity across craving-related ROIs | Yes, 2 visits |
| **Alcohol** | | | | | | |
| Karch et al. (2015) | Anterior ACC, insula, dorsolateral PFC | Continuous nf via veridical thermometer; 4 nf training runs, decrease trials using cognitive strategies during alcohol cues; pre- and post-nf resting state scans, fc | N = 34, 18-60 years old; AUD patients group n = 13, healthy controls n = 14, AUD sham nf group n = 2, ages, 43 (male) and 54 (female), control sham nf group n = 5, 22-43 years old, 20% female | Healthy control nf group, AUD sham nf group, healthy control sham nf group; neutral image condition | AUD nf group decreased ROI brain activation across nf training and increased connectivity between frontal cortex and subcortical areas | None |
| Kirsch et al. (2016) | VS and PFC | Continuous nf via veridical thermometer; monetary reward task as a functional localizer, 3 nf training runs, decrease trials using cognitive strategies and alcohol cues, post-nf transfer block | N = 38 heavy drinking college students; nf group n = 13, M age = 24.2, 38% female, yoke nf group n = 13, M age = 23.62, 15% female, non nf group n = 12, M age = 24.5, 17% female | Yoke nf group; no-nf group | True nf group decreased VS activation; neurofeedback group had significant prefrontal activation with correlation to decreased VA activation | None |
| **Cocaine** | | | | | | |
| Kirschner et al. (2018) | VTA/SN | Continuous nf via veridical moving ball; 2 nf training runs, increase trials using cognitions/mental imagery; pre- and post-nf runs | N = 50, 18-52 years old; cocaine use disorder group n = 22, M age = 29.7, 36% female, healthy controls n = 28, M age = 29.2, 50% female | Matched healthy control group; rest condition | Both groups increased VTA/SN activation; obsessive-compulsive cocaine users had weaker ability to modulate VTA/SN | None |
| Study | Target ROI(s)/Circuitry | Study Design | Participants | Control Condition/Group | Key Findings | Follow-Up |
|-------|-------------------------|--------------|--------------|--------------------------|--------------|-----------|
| Greer et al. (2014) | NAcc | Continuous nf via veridical thermometer; 2 nf training runs, increase and decrease trials using cognitions/mental imagery; pre- and post-nf runs; fc | N = 25, 20–40 years old; 44% female | Non-nf run condition | Better able to increase NAcc activation; nf improved up- and down-regulation; fc between NAcc and mPFC during nf | None |
| MacInnes et al. (2016) | VTA | Continuous nf via veridical thermometer; 3 nf training runs, increase trials using cognitions/mental imagery; pre- and post-nf runs; fc | N = 73; VTA nf group n = 19, M age = 24, 47% female, non-nf group n = 20, M age = 23, 50% female, NAcc nf group n = 20, M age = 23, 70% female, false nf group, n = 14, M age = 21, 64% female | 3 control groups: non-nf, NAcc nf, false nf; backward counting condition | VTA nf group increased ROI activation; fc increased in mesolimbic network after VTA nf | None |
| Sulzer et al. (2013b) | VTA/SN | Continuous nf via veridical moving ball; 2 nf training runs, increase trials using cognitions/mental imagery; pre- and post-nf runs; fc | N = 32, 24–35 years old, all male; VTA/SN nf group n = 15, inverse nf group n = 17 | Inverse nf group | VTA/SN nf group increased ROI activation; fc increased in nigrostriatal pathway after VTA/ SN nf | None |
| Emotion regulation | | | | | | |
| Caria et al. (2010) | Left anterior insula | Continuous nf via veridical thermometer; 4 nf runs, increase blocks using mental imagery, rated emotional valance and arousal from aversive images | N = 27, 23–40 years old (M age = 27.5), 56% female; nf group n = 9, 56% female, M age = 28, sham nf group n = 9, 56% female, M age = 27 mental imagery group n = 9, 56% female, M age = 26 | Sham nf group; mental imagery group; neutral images condition | Insula nf group increased ROI activation and rated aversive images more negatively | None |
| Cohen Kadosh et al. (2016) | Bilateral insula, amygdala, PFC | Continuous nf via veridical thermometer; 4 nf runs, increase and decrease trials using cognitions/mental imagery; effective connectivity (Granger causality) | N = 19, 7–16 years old (M age = 11.6), 42% female | Decrease condition | Better able to increase insula activation; correlations between the amygdala, PFC, and insula during up-regulation | None |
| Hamilton et al. (2011) | Subgenual ACC | Continuous nf via red line (ACC signal) and black line (whole brain minus ACC activity); 2 nf training runs, decrease blocks using positive mood strategies, pre- and post-nf runs; fc | N = 17, 18–50 years old, all female; nf group n = 8, yoked nf group n = 9 | Yoked nf group; baseline condition | ACC nf group decreased ROI activation and decreased correlation with PCC activation during nf training | None |
| Kouh et al. (2017) | Amygdala, dorsomedial PFC | Intermittent nf via a red number for successful trials and blue for non-successful trials; 3 nf runs, increase connectivity using cognitions/mental imagery, pre- and post-nf runs; effective connectivity (DCM) | N = 15, M age = 26.2, 53% female; nf group = 9, M age = 26.4, 56% female, sham nf group n = 6, M age = 25.7, 50% female | Sham nf group; neutral images condition | Increased connectivity between medial PFC and amygdala, corresponding to increased valance ratings | Yes, 3 visits |
| Lawrence et al. (2014) | Right anterior insula | Continuous nf via veridical thermometer; 4 nf training runs, increase blocks using cognitions/mental imagery, rated emotional valance and arousal from aversive images; pre- and post-nf runs; effective connectivity (DCM) | N = 24, 22–32 years old; nf group n = 16, 50% female, sham nf group n = 8, 50% female | Sham nf group; baseline and neutral images conditions | Insula nf group increased ROI activation; no transfer effects in arousal and valence ratings; reward learning mediated by dorsal ACC | None |
| Zotev et al. (2011) | Left amygdala | Continuous nf via text and color icons; 3 nf training runs, increase blocks using positive memories, pre- and post-nf runs; fc | N = 28, M age = 28.0, all male; nf group n = 14, M age = 27.5, sham nf group n = 14, M age = 28.4 | Sham nf group; count condition | Amygdala nf group increased ROI activation, increased fc across training and transfer runs | None |

**Note.** AUD = alcohol use disorder; fc = functional connectivity; DCM = dynamic causal modeling; nf = neurofeedback; ROI = region of interest; ACC = anterior cingulate cortex; NAcc = nucleus accumbens; PFC = prefrontal cortex; PPC = posterior cingulate cortex; SN = substantia nigra; VS = ventral striatum, VTA = ventral tegmental area.
instructed participants to use positive thoughts and recollections of prior events to increase activation in a targeted brain region involved with emotional processing. However, Caria et al. (2010) targeted brain activation in the anterior insula, a region involved in interoceptive processing that has projections both to and from the amygdala (Paulus and Stewart, 2014). In addition to the true neurofeedback group that received feedback linked to insula activation, one control group received sham neurofeedback from a different brain region, and another control group was asked to use the same mental strategies but had no neurofeedback. Findings demonstrated that the true neurofeedback group, in comparison to both control groups, was better able to modulate insula activation and reported significant changes in subjective emotional responses to image cues. Caria et al. (2010) identified learning-related brain activation not only in the right insula but also in the dorsolateral PFC, PCC, left inferior frontal gyrus, and left superior temporal gyrus as a result of neurofeedback training. These results show that even though self-regulation of emotional processing targeted to a single brain region can be learned through neurofeedback, neurofeedback training may also contribute to secondary effects within other emotion-related brain regions. In other words, neurofeedback may impact entire brain networks related to affective responding.

Targeting the right anterior insula (RAI) activation, Lawrence et al. (2014) used rtfMRI-nf to examine reinforcement learning associated with affective responding. The insula has previously been shown to be involved in emotion regulation related to substance use (Wilcox et al., 2016). A sample of healthy young adults were instructed to increase RAI activation using suggested strategies, such as recalling pleasant memories. Results indicated that the RAI neurofeedback group was better able to self-regulate RAI activation compared to the control group that received sham feedback from another region of the brain. However, there were no group differences in behavioral measures of arousal, valence ratings, or skin conductance response. Whole-brain analyses of brain activation during rtfMRI-nf indicated that the dorsal ACC mediated reward-related learning during RAI neurofeedback. Thus, decision-making associated with dorsal ACC activation may play a role in affective processing. This decision-making may translate to antecedents of substance use behavior, such as choosing to consume substances in an effort to enhance positive emotions or decrease negative emotions.

Deficits in emotion regulation that underly substance use behavior may be especially relevant to adolescents, due to the maturational delay in prefrontal brain systems involved in cognitive control that manage socioemotional reward responding (Casey, 2015; Shulman et al., 2016; Steinberg, 2010). Cohen Kadosh et al. (2016) examined if rtfMRI-nf could be used in a sample of children and adolescents (ranging in age from 7 to 16 years old, mean age of 11.6 years old) to up- and down-regulate bilateral insula activation, given the importance of this region in emotion regulation. Participants completed neurofeedback sessions in which they were instructed to think of positive thoughts to increase insula activation and remain as neutral as possible in their thoughts to decrease insula activation, with brain activation in the insula linked to a veridical thermometer image. Results showed that participants were able to successfully increase insula activation during neurofeedback. These findings did not significantly differ by sex or age of the sample. Effective connectivity analyses, measured by Granger information flow between brain regions activated during up- versus down-regulation tasks, indicated that neurofeedback was involved in the recruitment of other regions within the emotional regulation network, including the amygdala. They also found that prefrontal regions were not recruited during neurofeedback, positing that this finding may be attributable to the wide age range of the sample. Perhaps the greatest contribution of this work to the rtfMRI-nf literature is demonstrating the feasibility of using rtfMRI-nf in both child and adolescent samples.

### 4. Future directions for real-time fMRI neurofeedback research

Many existing rtfMRI-nf studies, including those in this review (Table 1), are described as “proof-of-concept” (e.g., Cohen Kadosh et al., 2016; Zotev et al., 2011) or “preliminary” (e.g., Hartwell et al., 2016; Li et al., 2013; Sulzer et al., 2013b) due to the novelty of the study design and often small sample sizes. Before rtfMRI-nf can be validated as an effective tool to elicit behavioral changes in substance using populations, or other clinical populations, follow-up work by these groups and additional studies is needed. Our review focused on discussing: 1) studies that used substance-using samples to examine rtfMRI-nf as a clinical tool to modulate craving-related brain activation; and 2) basic-science rtfMRI-nf studies that investigated the extent to which individuals could exert volitional control over brain function associated with addiction (i.e., reward responsivity and emotional regulation). These studies predominately used ROIs as targets for neurofeedback training, with a lesser extent of work using connectivity targets for neuromodulation (Table 2). Prior reviews have given detailed guidelines to improve the methodological rigor and clinical application of rtfMRI-nf studies (e.g., Fede et al., 2020; Stoeckel et al., 2014; Sulzer et al., 2013a; Thibault et al., 2018). Thus, the sections below are not exhaustive. They do, however, signify key issues and suggest recommendations for future rtfMRI-nf studies to address, especially those associated with substance use.

### 4.1. General design considerations

Fig. 1 outlines the typical real-time fMRI neurofeedback process, including selection of cognitive strategies for self-regulation of brain activity, calculation and display of the neurofeedback display to the subject, and analysis of rtfMRI-nf results. We will briefly review issues in each of these areas to consider in the experimental design of rtfMRI-nf studies.

#### 4.1.1. Self-regulation strategies

When designing a strategy for subjects to learn to modulate their own brain activity, one of the important decisions to be made is whether to have subjects use implicit or explicit cognitive strategies to achieve the desired response. Explicit strategies may allow for nominal homogeneity across subjects, afford the ability to rehearse without feedback, and use of already successfully established coping mechanisms (e.g., Emmert et al., 2017). However, cognitive strategies cannot be absolutely verified, and some brain regions or pathologies may not suggest an a priori approach (Sulzer et al., 2013b). Work with implicit strategies have shown that neurofeedback may be used to select the best regulation strategy (e.g., Caria et al. 2010, Lawrence et al. 2014), and spontaneous strategies have been demonstrated for control of brain activity (e.g., Kober et al., 2013; Shibata et al., 2011). Both strategy

---

**Table 2**

| Extraction Method       | Number of Studies | Study Reference               |
|-------------------------|-------------------|-------------------------------|
| Region of Interest      | 14                | Canterbury et al. (2013)      |
|                        |                   | Caria et al. (2010)           |
|                        |                   | Greer et al. (2014)           |
|                        |                   | Hamilton et al. (2011)        |
|                        |                   | Hanlon et al. (2013)          |
|                        |                   | Hartwell et al. (2016)        |
|                        |                   | Karch et al. (2015)           |
|                        |                   | Kirsch et al. (2016)          |
|                        |                   | Kirschner et al. (2018)       |
|                        |                   | Lawrence et al. (2014)        |
|                        |                   | Li et al. (2013)              |
|                        |                   | MacInnes et al. (2016)        |
|                        |                   | Sulzer et al. (2013b)         |
|                        |                   | Zotev et al., 2015            |
| Functional Connectivity | 1                | Kim et al. (2015)             |
| Effective Connectivity  | 2                | Cohen Kadosh et al. (2016)    |
|                        |                   | Koush et al., 2017           |
types may be advantageous in individuals with substance use disorders. Such individuals may benefit from explicit strategies to modulate brain function related to their use behavior if previously used strategies prior to rtfMRI-nf training were not successful (e.g. did not reduce cravings). Thus, explicit strategies for self-control may be useful when substance use decision-making occurs in real-life contexts. On the other hand, implicit strategies may be more individualized to participants and reduce preconceived beliefs of the researcher about which strategies are most effective for neuromodulation. Although compelling thoughts of substance use may impede the formation of implicit strategies, this strategy type may be ideal for individuals who have difficulty formulating explicit strategies. For both explicit and implicit strategies, careful documentation of individuals' employed strategies should be recorded after the rtfMRI-nf session (Ros et al., 2019). This information may aid in the development of more targeted interventions. Further discussion on self-regulation strategies can be found in the reviews of Shibata et al. (2011) and Sulzer et al. (2013a).

4.1.2. Feedback signal calculation and display

The specific brain activation used for the neurofeedback signal plays a central role in the success of any particular rtfMRI-nf study. As displayed in the studies reviewed above, a majority of studies use a specific ROI or set of ROIs, based on knowledge of the pathology and/or modulation targets. Although this has been proven successful in a number of studies, it is worthwhile to acknowledge that multiple brain regions and networks are involved in any fMRI task, especially one as complex as rtfMRI-nf, where the base task of interest (e.g. craving) also now has the addition of attentional and executive control networks used in the regulation activity. In addition, any particular brain region varies in anatomy and functional response across subjects, control regions may vary as a function of age (Johnson et al., 2015), and using prescribed regions may prevent finding results from implicit strategies (Lange et al., 1999). Considering network or full brain activations, and employing multivariate approaches (e.g. LaConte et al., 2007), may better capture individual response in fMRI neurofeedback studies (Paret et al., 2019). See LaConte (2011) and Young et al. (2017) for more in-depth review of various approaches.

In terms of feedback display to the subject, choices need to be made in terms of complexity and update rate. A majority of rtfMRI-nf studies use some version of a continually updating thermometer display to relay brain activity information to the subjects (Fedele et al., 2020). However, too simple a display can lead to subject fatigue (Paret et al., 2019), and more social feedback may be more engaging for subjects (Mathiak et al., 2015). In terms of timing, studies have found that intermittent feedback instead of continuous feedback may be more effective in helping subjects achieve self-regulation (Hellrung et al., 2018; Johnson et al., 2012). An additional advantage of intermittent feedback is that time restriction is reduced, thereby allowing for more complex data processing (Scheinost et al., 2020). The choice between continuous/intermittent feedback may depend on the self-regulation process being employed (Oblak et al., 2017). Impacting these decisions is the fact that the act of receiving neurofeedback information may interfere with the desired cognitive control outcome (Lubianiker et al., 2019).

4.1.3. Neurofeedback control conditions

As with any potential therapy, rtfMRI-nf needs to have appropriate control conditions in order to demonstrate that observed changes are due to the proposed intervention. The most common control conditions are using runs with sham neurofeedback to control for the subject expectancy (placebo) effect, while runs with no feedback control for the neurofeedback process. Both sham and no feedback methods should be used in tandem to examine the actual benefits of neurofeedback (Fedele et al., 2020). It is also possible to have a discrepancy between subject expected outcome and observed feedback measure, which can result in worsened expectations (“nocebo effect”, Zubieta and Stohler, 2009). Other control condition possibilities include the inverted response condition (mentioned above for the Sulzer et al. (2013a) study), and using positive control conditions, in which the observed rtfMRI-nf modulation of activity is compared to the “gold-standard” task method for eliciting activity in the given area or network (Berman et al., 2011, Sulzer et al. (2013a). See Sorger et al. (2019) for further discussion.

As a relatively recent neuroimaging methodology, the clinical applications of rtfMRI-nf are not yet universally validated. However, investigators using rtfMRI-nf are beginning to employ more robust study designs and control conditions. Although focused on patients with depression, a recently published randomized clinical trial represents an exemplar approach to measuring the therapeutic potential of rtfMRI-nf. In their double-blind, placebo-controlled study, Young et al. (2017) found that neurofeedback training aimed at increasing amygdala response to positive autobiographical memory recall significantly decreased depressive symptoms. Randomized controlled trials are also warranted in rtfMRI-nf interventions for individuals with substance use disorders.

4.2. Sample selection

In a few of the studies included in this review, individual variability in the ability to modulate brain activation through rtfMRI-nf was reported (e.g., Karch et al., 2015). This may be due, in part, to the ages of participants included in these studies. Most of the studies we reviewed used young to middle adult samples. Others used much wider age ranges of participants (Hanlon et al., 2013; Li et al., 2013), with some studies including participants ages 18–50 years old (Hamilton et al., 2011) and 18 to 60 years old (Canteberry et al., 2013; Hartwell et al., 2016; Karch et al., 2015). Other work beyond the focus of our review has indeed identified age differences in neurofeedback performance, such as a negative correlation between age and default mode network neurofeedback performance in a non-pathological sample of adults aged 20–45 years old (Skounis & Scharnowski, 2019).

Age is relevant both to rtfMRI-nf studies with substance use samples and basic science research pertinent to addiction, particularly the distinction between adolescence, young adulthood, and older adulthood. Adolescence is when substance use is most likely to begin; according to the results from the 2018 National Survey on Drug Use and Health (Substance Abuse and Mental Health Services Administration, 2019), approximately 2.2 million adolescents between the ages of 12 and 17 drank alcohol in the past month. About 1 in 6 adolescents within this age span were past-year illicit drug users. Substance use then tends to peak in young adulthood, with annual prevalence rates of illicit drug use (including marijuana) peaking between the ages of 21 and 24 at 47% and past two-week binge drinking most frequent among 23- and 24-year-olds at 37% (Schulenberg et al., 2019). These rates are concerning, considering that the neural structure and function of youths' brains are particularly sensitive to the neurotoxic effects of substance use (Jacobs & Tapert, 2013; Lisdahl et al., 2013; Squeglia & Gray, 2016). Neural sensitivities are coupled with aspects of brain development—specifically the protracted maturation of brain systems involved in self-regulation of reward and emotional responding—that make adolescents and young adults especially susceptible to substance use (Casey, 2015; Ernst, 2014; Shulman et al., 2016; Steinberg, 2010). Thus, adolescents and young adults should be viewed as a unique study population in terms of both substance use behavior and self-regulation.

Heightened levels of substance use and deficits in self-regulation are not limited to youth. For example, the 2018 Monitoring the Future survey found that among individuals 19–60 years old, daily drinking peaked at age 60 (12%) and nearly a fifth of 50-year-old respondents reported current binge drinking (Schulenberg et al., 2019). The baby-boomer generation is most likely to report the use of alcohol, tobacco, marijuana, and the misuse of prescription drugs as they age into older adulthood (Schulte & Hser, 2013). Deficits in self-regulation attributed to the effects of prolonged substance use may potentially interfere with
older adults’ ability to exert volitional control over brain activation during rtfMRI-nf. Furthermore, the efficiency of brain systems involved in inhibitory control shows age-related declines (Coxon et al., 2016). To better account for potential age effects, it would be beneficial for future rtfMRI-nf studies to test their neurofeedback paradigm separately by age groups, or at the minimum control for age in statistical analyses that compare group differences in the ability to modulate brain activation.

Variation in substance use patterns has been found not only in relation to age, but also by other sociodemographic factors, such as sex. Although the gender gap for rates of substance use is narrowing (Keyes et al., 2010), men often have higher levels of drug and alcohol use than women (Schulenberg et al., 2019). In addition, pathways to substance use typically differ by sex. Women are more likely to engage in substance use as a way to cope with depression and anxiety, and this internalizing pathway to substance use involves deficits in emotion regulation and heightened amygdala activation (Hardee et al., 2018). Men, on the other hand, are more likely to engage in substance use through an externalizing pathway characterized by impulsivity and aggression (Hicks et al., 2007). The externalizing pathway predominately involves brain activation associated with reward responsibility. Furthermore, Becker & Koob (2016) describe sex differences in each stage of addiction; women tend to experience intoxication at lower levels of consumption, report greater deficits in emotional regulation during withdrawal, and are more prone to stress-induced relapse compared to men. Despite these sex-related differences, few rtfMRI-nf studies have investigated if the ability to modulate brain activation through neurofeedback differs between men and women. Doing so is especially important for rtfMRI-nf research that targets brain activation in affective and reward systems, due to their association with internalizing and externalizing behaviors, respectively, that underlie substance use behavior. In sum, identifying brain systems that differ between certain study samples prior to neurofeedback sessions, as well as group differences in neurofeedback performance, has important implications for improving the efficacy and generalizability of both clinical and basic science rtfMRI-nf research.

4.3. rtfMRI-nf as a treatment modality

Relapse following substance use treatment is common (e.g., Hendershot et al., 2011), with 40–60% of patients who seek treatment for drug addiction reporting relapse (National Institute on Drug Abuse, 2018). Inability to manage cravings and deficits in negative affect have been identified as predictors of continued substance use (e.g., McKay, 2011; Plassceki, 2006). Thus, developing evidence-based substance use interventions that lead to sustained reductions in drug and alcohol use is an important goal of clinical substance use research. Behavioral therapies, particularly CBT, are the most commonly used type of addiction treatment and are also sometimes used in conjunction with medications (e.g., nicotine replacement, naltrexone). Success rates of current addiction treatments are modest but can lead to sustained reductions or elimination of problem substance use. According to NIDA’s recent publication on research-based addiction treatment, successful outcomes are contingent upon the level of addiction severity, the adequacy of match between treatment services and patient needs and characteristics, and treatment plan adjustments based on continual evaluation and modification (National Institute on Drug Abuse, 2018).

Perhaps one of the most appealing aspects of using rtfMRI-nf as a treatment for substance use disorders is that it directly taps into both brain function and behavior involved in addiction. However, for rtfMRI-nf to become a viable treatment for addiction, the transferability of learning effects acquired through neurofeedback into various contexts and over time needs to be determined. An essential facet of rtfMRI-nf is operant conditioning-based learning based on reinforcement (i.e., receiving neurofeedback linked to brain activation). Whether this learning extends into various contexts beyond the laboratory setting has been a potential limitation, and area of continued research, for rtfMRI-nf studies (Thibault et al., 2018). Studies that include transfer runs provide an indicator of learning effects (e.g., Greer et al., 2014; Zotev et al., 2011). All of the papers studying substance use samples (nicotine, alcohol, and cocaine) discussed in this review measured subjective craving before and after rtfMRI-nf tasks with the exception of Kirschner et al. (2018). Assessing sustained learning effects to reduce craving is essential in understanding the potential utility of rtfMRI-nf as a viable substance use treatment. Yet, a large majority of existing rtfMRI-nf studies only report changes in self-regulation immediately following the neurofeedback session and do not assess continued learning effects. In our review, only 5 out of 17 studies had any follow-up assessments. Identifying the potential long-term effects of neurofeedback training on sustained reductions in substance use is essential in order for rtfMRI-nf to be used as a treatment modality for addiction. Emerging evidence suggests that rtfMRI-nf can indeed have lasting therapeutic effects on behavioral modification and symptom reduction (Mehler et al., 2018; Rance et al., 2018). Continued research that incorporates follow-up assessments on rtfMRI-nf learning effects is needed to validate the efficacy of neurofeedback training.

Conducting rtfMRI-nf scans is expensive and requires extensive training. As such, research is needed to determine the ideal number of rtfMRI-nf sessions required to produce reductions in both craving and substance use behavior. For example, the only two studies that tested neurofeedback training to reduce craving-related brain activation in alcohol-using samples were conducted in a single visit (Karch et al., 2015; Kirsch et al., 2016). Although Canterberry et al. (2013), Hanlon et al. (2013), Hartwell et al. (2016), and Kim et al. (2015) used a multi-visit study design in their rtfMRI-nf studies of nicotine users, none of these studies conducted follow-up assessments to determine the extent to which neurofeedback reduced nicotine craving over time and beyond the laboratory context. These preliminary studies were novel in their approach, and therefore framed in relation to the feasibility of implementing their study paradigm. Additional research is needed to determine what level of neurofeedback training impacts substance use. Consistent with our prior discussion on age and sex differences, this work would also benefit from investigating potential sociodemographic differences in the level of neurofeedback training needed to produce significant reductions in substance use. A preprint of a recent meta-analysis of pre-training effects on neurofeedback learning indicates that there is much individual variability in the ability to modulate brain activity through neurofeedback, even if participants received neurofeedback training (Haugg et al., 2020). The authors suggest that variability in learning curves to successfully regulate brain function through neurofeedback may be attributable to individual differences in cognitive capacities, especially for clinical populations with pre-existing deficits in self-regulation. Thus, pre-training levels of brain function is an important aspect to consider for future rtfMRI-nf research, especially work in individuals with substance use disorders.

The practicality of using rtfMRI-nf as a treatment modality, and its role in informing more individualized substance use interventions, would also likely benefit from expanding rtfMRI-nf research to a wider range of substances. As described in this review, substance use-related rtfMRI-nf studies have focused predominately on nicotine. At this time, however, marijuana is the most commonly used illicit drug in the United States (Schulenberg et al., 2019), and an increasing number of states within the United States are legalizing the recreational use of marijuana. Therefore, rtfMRI-nf studies of marijuana-using samples are warranted. This is particularly important given research findings indicating the potential neurocognitive impacts of marijuana use. For example, Martz et al. (2016) found that marijuana use may weaken the brain’s response to natural rewards over time. These findings held even after accounting for participants’ alcohol and tobacco use, suggesting that marijuana may have a unique impact on anticipatory reward sensitivity.

Regardless of the feasibility of using rtfMRI-nf for widespread and
readily available clinical treatment, it has benefits as a research tool to inform clinical practice. For example, in relation to addiction, rtfMRI-nf can be developed as a tool for quantifying individuals’ capacity for volitional control over socioemotional, reward responding, allowing researchers to more directly test hypotheses about the role of such capacities in substance use risk. This work also has the potential to more precisely target brain activity involved in addiction, which may then be used to build more individualized interventions to strengthen self-control over substance use behavior.

5. Conclusions

This review identified key applications of rtfMRI-nf relevant to addiction. First, we presented findings demonstrating the potential of rtfMRI-nf to train substance use populations to decrease craving-related brain activation. This training represents a facet of self-control at the neural level that may help individuals struggling with substance use problems develop self-efficacy in their ability to connect their thoughts to a visible marker of brain activation related to their substance use behavior (e.g., craving). In this way, strategies developed during rtfMRI-nf sessions may be used to strengthen self-control over drug and alcohol consumption. However, whether or not this training is sustained across varying contexts and over time remains largely unknown. Although this review provides the first overview of literature on rtfMRI-nf relevant to addiction, non-systematic reviews are limited in their precision of searching existing literature and may be susceptible to bias. As more rtfMRI-nf studies relevant to addiction emerge in the literature, so too should more rigorous reviews (e.g., systematic reviews, meta-analyses) in order to provide a more thorough qualitative and quantitative synthesis of existing data.

Second, we have also reviewed rtfMRI-nf research in healthy populations that targeted neuromodulation over brain activation associated with addiction. We specifically targeted studies assessing volitional control over reward responsivity or emotional regulation, two neural correlates of substance use that have been well-documented in the literature. Findings from this research demonstrated areas of the brain associated with addiction that participants can self-regulate through rtfMRI-nf. The extent to which specific brain regions can be up- and/or down-regulated helps better understand their function and establishes these regions as potential targets for clinical interventions.

Third, we provide recommendations for future rtfMRI-nf research relevant to addiction. This area of study is still in its infancy, and existing work is largely preliminary. However, we offer suggestions to improve the design of rtfMRI-nf studies, based on this limited body of research and other clinical applications of rtfMRI-nf. Our recommendations include post-scan documentation of participant strategies for neuromodulation, using multivariate approaches for neurofeedback training, using the type self-regulation processes to inform decisions to use continuous or intermittent feedback, conducting randomized controlled trials, careful sample selection (e.g., accounting for potential effects of age, sex, and sociodemographic factors on neuromodulation), and the importance of follow-up assessments in clinical studies to assess the impact of neurofeedback training on substance use behavior. In sum, there is much promise, and yet much work still to be done, applying rtfMRI-nf to the study and treatment of addiction.

Funding

During manuscript preparation, Dr. Martz was supported by the National Institute on Alcohol Abuse and Alcoholism of the National Institutes of Health under Award Numbers R01 AA027558 and Dr. Heitzeg was supported by R01 AA12217. Dr. Heitzeg was also supported by the National Institute on Drug Abuse of the National Institutes of Health under Award Number R01 DA027261. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

CRediT authorship contribution statement

Meghan E. Martz: Writing - review & editing, Conceptualization, Writing - original draft, Visualization. Tabatha Hart: Writing - review & editing, Visualization. Mary M. Heitzeg: . Scott J. Peltier: Writing - peer review & editing, Visualization.

References

Allman, J.M., Hakeem, A., Erwin, J.M., Nimchinsky, E., Hof, P., 2001. The anterior cingulate cortex: the evolution of an interface between emotion and cognition. Ann. N. Y. Acad. Sci. 935, 107–117.
Becker, J.B., Koob, G.F., 2016. Sex differences in animal models: Focus on addiction. Pharmacol. Rev. 68 (2), 242–263.
Berman, B.D., Horovitz, S.G., Venkataraman, G., Hallett, M., 2011. Self-modulation of primary motor cortex activity with motor and motor imagery tasks using real-time fMRI-based neurofeedback. NeuroImage 59 (2), 917–925.
Bush, G., Luo, P., Posner, M.I., 2000. Cognitive and emotional influences in anterior cingulate cortex, Trends in Cognit. Sci. 4 (6), 215–222.
Canterberry, M., Hanlon, C.A., Hartwell, K.J., Li, X., Owens, M., LeMatty, T., George, M.S., 2013. Sustained reduction of nicotine craving with real-time neurofeedback: exploring the role of severity of dependence. Nicotine Tob. Res. 15 (12), 2120–2124.
Caria, A., Sitaram, R., Veit, R., Beglomini, C., Birbaumer, N., 2010. Volitional control of anterior insula activity modulates the response to aversive stimuli. A real-time functional magnetic resonance imaging study. Biol. Psychiatry 68 (5), 425–432.
Carroll, K.M., Kilik, B.D., 2017. Cognitive behavioral interventions for alcohol and drug use disorders: Through the stage model and back again. Physiol. Behav. 31 (8), 547–561.
Casey, B.J., 2015. Beyond simple models of self-control to circuit-based accounts of adolescent behavior. Annu. Rev. Psychol. 66, 295–319.
Chung, T., Norouzi, A., Carroll, K.M., Potenza, M.N., Hutchison, K., Calhoun, V.D., Kober, H., 2016. Brain mechanisms of change in addiction treatment: Models, methods, and emerging findings. Curr. Addict. Rep. 3 (3), 332–342.
Cohen Kadosh, K., Luo, Q., de Burca, C., Sokunbi, M.O., Feng, J., Linden, D.E.J., Lau, J.Y.F., 2016. Using real-time fMRI to influence effective connectivity in the developing emotion regulation network. NeuroImage 125, 616–626.
Coxon, J.P., Golbe, D.J., Leunissen, I., Van Ampe, A., Wenderoth, N., Swinnen, S.P., 2016. Functional brain activation associated with inhibitory control deficits in older adults. Cereb. Cortex 26 (1), 12–22.
Ekhtiari, H., Faghiri, A., Oghlanian, M.A., Paulus, M.P., 2016. Functional neuroimaging for addiction medicine: from mechanisms to practical considerations. In: Progr. Brain Res. 224. Elsevier, pp. 129–153.
Emmert, K., Kopel, R., Roush, Y., Maire, R., Senn, P., Van De Ville, D., Haller, S., 2017. Continuous vs. intermittent neurofeedback to regulate auditory cortex activity of tinnitus patients using real-time fMRI pilot study. NeuroImage: Clinical 14, 97–104.
Ernst, M., 2014. The triadic model as a perspective for the study of adolescent motivation. Brain Cogn. 89, 104–111.
Fede, S.J., Dean, S.F., Manuweera, T., Momenan, R., 2020. A guide to literature informed decisions in the design of real time fMRI neurofeedback studies: a systematic review. Front. Hum. Neurosci. 14.
Greer, S.M., Trujillo, A.J., Glover, G.H., Knutson, B., 2014. Control of nucleus accumbens activity with neurofeedback. NeuroImage 96, 237–244.
Hamilton, J.P., Glover, G.H., Hsu, J.J., Johnson, R.F., Gotlib, I.H., 2011. Modulation of subgenual anterior cingulate cortex activity with real-time neurofeedback. Hum. Brain Mapp. 32 (11), 22–31.
Hampson, M., Driessen, N.R., Skudlarski, P., Gore, J.C., Constable, R.T., 2006. Brain connectivity related to working memory performance. J. Neurosci. 26, 13338–13343.
Hanlon, C.A., Hartwell, K.J., Canterberry, M., Li, X., Owens, M., LeMatty, T., George, M.S., 2013. Reduction of cue-induced craving through realtime neurofeedback in nicotine users: The role of region of interest selection and multiple visits. Psychiatry Res. Neuroimag. 213 (1), 79–81.
Hardee, J.E., Cope, I.M., Martz, M.E., Heitzeg, M.M., 2018. Review of neurobiological influences on externalizing and internalizing pathways to alcohol use disorder. Curr. Behav. Neurosci. Rep. 5 (4), 249–262.
Hartwell, K.J., Hanlon, C.A., Li, X., Borckardt, J.J., Canterberry, M., Prisciandaro, J.J., Sayeed, K.K., 2016. Individualized real-time fMRI neurofeedback to attenuate craving in nicotine-dependent smokers. J. Psychiatry Neurosci. 41 (1), 48–55.
Haug, A., Sladky, R., Skouras, S., McDonald, A., Craddock, C., Kirchner, M., …& Scharnowski, R., 2020. Can we predict real-time fMRI neurofeedback learning success from pre-training brain activity? bioRxiv. https://www.biorxiv.org/content/10.1101/2020.01.15.906388v1.full.
Hellriegl, L., Dietrich, A., Hollmann, M., Pfeifer, B., Kalberlah, C., Boggenhofer, E., Horstmann, A., 2018. Interim comparison to continuous real-time fMRI neurofeedback boosts control over amygdala activation. NeuroImage 166, 198–208.
Hendershot, C.S., Wiktiszewit, K., George, W.H., Marlatt, G.A., 2011. Relapse prevention for addictive behaviors. Subst. Abuse Treatm. Prevent. Policy 6 (1), 17.
Hicks, B.M., Blonigen, D.M., Kramer, F.D., Krueger, R.F., Patrick, C.J., Iacono, W.G., McGue, M., 2007. Gender differences and developmental change in externalizing disorders from late adolescence to early adulthood: A longitudinal twin study. J. Abnorm. Psychol. 116 (3), 433–447.
Hussong, A.M., Jones, D.J., Stein, G.L., Bascom, D.H., Boeding, S., 2011. An internalizing pathway to alcohol use and disorder. Psychol. Addict. Behav. 25 (3), 390–404.
