Moderate effects of noninvasive brain stimulation of the frontal cortex for improving negative symptoms in schizophrenia: Meta-analysis of controlled trials

André Aleman\textsuperscript{a,b,c,*}, Stefanie Enriquez-Geppert\textsuperscript{c}, Henderikus Knegtering\textsuperscript{b,d}, Jozarni J. Diabac-de Lange\textsuperscript{b}

\textsuperscript{a} Shenzhen Key Laboratory of Affective and Social Cognitive Science, Shenzhen University, Shenzhen, 518060, China
\textsuperscript{b} University of Groningen, University Medical Center Groningen, Department of Neuroscience, Groningen, The Netherlands
\textsuperscript{c} University of Groningen, Department of Clinical and Developmental Neuropsychology, Groningen, The Netherlands
\textsuperscript{d} Lentis Mental Health Center, Groningen, The Netherlands

**ABSTRACT**

**Background:** Negative symptoms in schizophrenia concern a clinically relevant reduction of goal-directed behavior that strongly and negatively impacts daily functioning. Existing treatments are of marginal effect and novel approaches are needed. Noninvasive neurostimulation by means of repetitive transcranial magnetic stimulation (rTMS) and transcranial direct current stimulation (tDCS) are novel approaches that may hold promise. **Objectives:** To provide a quantitative integration of the published evidence regarding effects of rTMS and tDCS over the frontal cortex on negative symptoms, including an analysis of effects of sham stimulation. **Methods:** Meta-analysis was applied, using a random effects model, to calculate mean weighted effect sizes (Cohen’s d). Heterogeneity was assessed by using Cochrans Q and I\(^2\) tests. **Results:** For rTMS treatment, the mean weighted effect size compared to sham stimulation was 0.64 (0.32–0.96; \(k = 22\), total \(N = 827\)). Studies with younger participants showed stronger effects as compared to studies with older participants. For tDCS studies a mean weighted effect size of 0.50 (−0.07 to 1.07; \(k = 5\), total \(N = 134\)) was found. For all frontal noninvasive neurostimulation studies together (i.e., TMS and tDCS studies combined) active stimulation was superior to sham, the mean weighted effect size was 0.61 (24 studies, 27 comparisons, 95% confidence interval 0.33–0.89; total \(N = 961\)). Sham rTMS (baseline - posttreatment comparison) showed a significant improvement of negative symptoms, \(d = 0.31\) (0.09–0.52; \(k = 16\), total \(N = 333\)). Whereas previous meta-analyses were underpowered, our meta-analysis had a power of 0.87 to detect a small effect. **Conclusions:** The available evidence indicates that noninvasive prefrontal neurostimulation can improve negative symptoms. This finding suggests a causal role for the lateral frontal cortex in self-initiated goal-directed behavior. The evidence is stronger for rTMS than for tDCS, although this may be due to the small number of studies as yet with tDCS. More research is needed to establish moderator variables that may affect response to neurostimulation and to optimize treatment parameters in order to achieve stable and durable (and thus clinically relevant) effects.

1. Introduction

Negative symptoms in schizophrenia concern a markedly reduced interest and initiative, manifested in reductions of goal-directed behavior. Such reductions are evident in symptoms such as social withdrawal, apathy, alogia, anhedonia and reduced emotional expression. High levels of negative symptoms are a hallmark of poor outcome in schizophrenia (Tek et al., 2001; Galderisi et al., 2013; Üçok and Ergül, 2014). Unfortunately, treatment effects of conventional approaches with antipsychotics, other pharmacological agents or psychosocial interventions are limited and not clinical significant when it comes to reducing negative symptoms and improving social outcome (Aleman et al., 2017; Arango et al., 2013; Fusar-Poli et al., 2015; Lincoln et al., 2011). Therefore, the development of novel approaches is of great importance (cf. Millan et al., 2014). Noninvasive brain stimulation offers a novel approach in the
treatment of negative symptoms (Aleman, 2013). Several studies have used repetitive transcranial magnetic stimulation (rTMS) to enhance activation of the frontal cortex in patients with schizophrenia. Five previous meta-analyses (see Table 1) have synthesized evidence published up to 2014 and the first three found small to medium average effect sizes which were statistically significant and favoured rTMS over placebo stimulation (Freitas et al., 2009; Dlabac-de Lange et al., 2010; Shi et al., 2014). The meta-analysis by Fusar-Poli et al. (2015) included a total of eight studies published before December 2013 (with a total of 177 patients). That meta-analysis reported a mean weighted effect size of 0.23, statistically nonsignificant. Unfortunately, several published trials were not included and the effect size of one trial (Fitzgerald et al., 2008) was erroneously included as favouring sham stimulation, whereas the published data favoured active stimulation. That is, the article reported a reduction of negative symptoms of 16.7 points on the SANS (Schedule for Assessment of Negative Symptoms) in the rTMS group, and a reduction of only 6.8 points in the sham group. We identified ten recently published studies (that were not included in the last meta-analysis) and therefore an updated meta-analysis would be timely.

Recently, methods of noninvasive brain stimulation other than rTMS have been employed to improve negative symptoms. Specifically, transcranial direct current stimulation (tDCS), was used in several studies. Brain stimulation with tDCS involves weak electric fields, with currents of 1–2 mA. Precise mechanisms of action remain to be fully elucidated, but it is known that tDCS does not induce neuronal firing by supra-threshold neuronal membrane depolarization, as happens in rTMS, but rather modulates spontaneous neuronal network activity. This occurs through a tDCS polarity-dependent shift (polarization) of resting membrane potential (Priori et al., 2009; Paulus, 2011). Cortical activity and excitability may be enhanced through anodal tDCS stimulation, whereas cathodal tDCS stimulation may reduce excitability. rTMS and tDCS are non-invasive brain stimulation methods that can be used without anaesthesia (unlike electroconvulsive therapy, ECT) and have been used for experimental treatment of negative symptoms. Although they may well differ in their mechanism of action, rTMS and tDCS share a favorable side-effects profile. We chose to review both methods together as they both have been used to address the question of targeting prefrontal excitability to improve negative symptoms and are of similar interest to clinicians.

Studies using noninvasive neurostimulation to improve negative symptoms in schizophrenia have typically targeted the prefrontal cortex, more specifically the dorsolateral prefrontal cortex (DLPFC). This is based on neuroimaging findings of reduced DLPFC activation in patients with negative symptoms (e.g., Wolkin et al., 1992). Thus, the aim of the treatment is to increase excitability of the DLPFC. It should be noted that the DLPFC has a central role in functional neuroanatomical models of goal-directed behavior (Aarts et al., 2011; Yamagata et al., 2012). Although many details remain to be elucidated regarding the precise role of different areas and their connections, the DLPFC can be considered to be a key hub in a frontostriatal network (that may also involve premotor cortex and thalamus) subserving action planning, selection, preparation and evaluation. Neurostimulation studies can contribute to establishing a causal role for the DLPFC in goal-directed behavior.

We here integrate the published evidence regarding effects of non-invasive neurostimulation over the frontal cortex on negative symptoms using meta-analysis. Besides computing the mean weighted effect of rTMS versus sham stimulation across studies, we also estimated the effect of sham stimulation alone to estimate the magnitude of the placebo effect. Moreover, we present several additional analyses to identify potential moderators of the effect of brain stimulation.

### Table 1
Comparison of current meta-analysis with previously published meta-analyses of noninvasive brain stimulation for treatment of negative symptoms. Power to detect small effect sizes was computed with software provided online by AI-Therapy Statistics (https://www.ai-therapy.com/psychology-statistics/power-calculator).

<table>
<thead>
<tr>
<th>Study</th>
<th>Date range</th>
<th>Trials</th>
<th>N subjects</th>
<th>Power to detect small ES (Cohen d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freitas et al. (2009)</td>
<td>1999–2007</td>
<td>8</td>
<td>107</td>
<td>0.18</td>
</tr>
<tr>
<td>Dlabac-de Lange et al. (2010)</td>
<td>1999–2008</td>
<td>9</td>
<td>213</td>
<td>0.31</td>
</tr>
<tr>
<td>Slotten et al. (2010)</td>
<td>1999–2008</td>
<td>7</td>
<td>148</td>
<td>0.23</td>
</tr>
<tr>
<td>Shi et al. (2014)</td>
<td>1999–2013</td>
<td>16</td>
<td>348</td>
<td>0.46</td>
</tr>
<tr>
<td>Fusar-Poli et al. (2015)</td>
<td>1999–2013</td>
<td>8</td>
<td>177</td>
<td>0.26</td>
</tr>
<tr>
<td>He et al. (2017)</td>
<td>1999–2015</td>
<td>7</td>
<td>390</td>
<td>0.50</td>
</tr>
<tr>
<td>Current meta-analysis</td>
<td>1999–2017 (2.5 additional years)</td>
<td>24 (19 + 5 tDCS; 50% increase)</td>
<td>961 (147% increase)</td>
<td>0.87 (74% increase)</td>
</tr>
</tbody>
</table>

* The power to detect a small effect size of 0.2 (cf. Cohen, 1988).

2. Methods

2.1. Literature search and study selection

We included studies published up to December 2017. Studies were identified initially by performing a literature search in PubMed through June 2016 and by conducting a cross-reference search of the eligible articles to identify additional studies not found in the electronic search. The search terms used were “transcranial magnetic stimulation”, “transcranial direct current”, and “negative symptoms”. We also conducted additional searches in Web of Science (Thomson Reuters) up to December 2017 to make sure we did not miss studies. Web of Science includes Social and Behavioral Sciences in addition to Medical Sciences. This additional search did not yield previously unidentified studies. The primary outcome measure was reduction of negative symptoms as measured with the Brief Psychiatric Rating Scale (BPRS), the Scale for the Assessment of Negative Symptoms (SANS), or the negative symptom subscale of the Positive and Negative Syndrome Scale (PANSS). Criteria for inclusion in the meta-analysis were a parallel or crossover design with sham control in patients with schizophrenia, schizoaffective disorder, or schizotypal disorder. Crossover trials with a wash-out phase of less than 4 weeks were excluded (Dlabac-de Lange et al., 2010). Only studies using rTMS of the prefrontal cortex, which is the focus of the vast majority of studies and of this review, were included. If there was insufficient information in the article to calculate the effect size, the corresponding author was contacted. In case no sufficient data for calculation of effect sizes could be obtained from article or authors, studies were excluded from the meta-analysis.

2.2. Statistical analysis

Individual effect sizes (Cohen d) of each study were calculated using the effect size program developed by Wilson (cf. http://www.campbellcollaboration.org/escalc). Whenever possible we computed standardized mean gain effect sizes (cf. Lipsey and Wilson, 2001), to account for the fact that the same sample is measured twice (pre-post contrast). When no pre- and post means and SDs were given for each group, but sufficient statistical information in the form of mean change (and SD), or precise t, F, or p-values was available, the standardized
difference (d) was computed using the same software. When data on different scales rating the same effect were available, the data were pooled, calculating a standardized mean difference. If only means but no standard deviations were reported, we used the mean standard deviation of all the other studies as an estimate (this procedure was necessary for only one study, Schneider et al., 2008). A random effects model was used, and the mean weighted effect size was calculated by using Review Manager 5.0, developed by The Cochrane Collaboration. Individual effect sizes were weighted by the standard error of the estimate. Heterogeneity refers to variability among studies, which may be caused by clinical and methodological diversity. Significant heterogeneity limits a reliable, unequivocal interpretation of the results. Heterogeneity was assessed by using Cochrans Q and I² tests. Cochran’s Q is calculated as the weighted sum of squared differences between individual study effects and the pooled effect across studies, with the weights being those used in the pooling method. Q is distributed as a chi-square statistic with k (number of studies) minus 1 degrees of freedom. The I² statistic describes the percentage of variation across studies that is due to heterogeneity rather than chance. For more information regarding these measures, we refer to meta-analysis handbooks (e.g. Lipsey and Wilson, 2001; Borenstein et al., 2009).

3. Results

The search yielded 90 publications (77 for the combination with rTMS and 13 for TDCS). Of these, 66 were excluded because they did not fulfill the inclusion criteria (see Fig. 1). The remaining articles contained 24 studies (27 independent comparisons) reporting on the difference between active and sham stimulation (total N of 966 patients) that could be included for meta-analysis (some articles contained more than one independent comparison, see Fig. 2). Compared to the largest previous meta-analysis (see Table 1), our meta-analysis contained data of 64% more patients and represents a 47% increase in power.

3.1. rTMS studies only

Information regarding the included studies applying rTMS is given in Table 2. For only rTMS treatment, the mean weighted effect size was 0.64 (0.32–0.96; I² = 79%, k = 22, total N = 825), with a stronger improvement for active stimulation as compared to sham. The study by Goyal et al. (2007) and the study (with four experimental groups) by Zhao et al. (2014) showed much larger effect sizes than the other studies, and could be considered to be statistically outliers. We therefore conducted an analysis without these studies, to see if a significant effect of rTMS would survive. Without Goyal et al. (2007) and Zhao et al. (2014) the mean weighted effect size became 0.31 (0.12–0.50; I² = 30%, k = 18, total N = 721). Heterogeneity was nonsignificant for the latter analysis, Q(17) = 24.40, p = 0.11 (Fig. 3).

3.2. Potential moderators of effect

We conducted several analyses separately for studies grouped according to a relevant variable that could affect effect size. Again, the outlier studies (Goyal et al., 2007 and Zhao et al., 2014) were not included, so as not to bias the results. When studies using a frequency of

---

![Flowchart](image-url)
Fig. 2. Forest plot of effect sizes for active versus sham rTMS treatment over the DLPFC for improving negative symptoms.

Table 2
Studies included in the meta-analysis applying rTMS.

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>N</th>
<th>Location</th>
<th>rTMS frequency</th>
<th>rTMS intensity</th>
<th>number of stimuli</th>
<th>duration, days</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barr et al. (2012)</td>
<td>25</td>
<td>bilateral DLPFC</td>
<td>20</td>
<td>90% MT</td>
<td>30000</td>
<td>20</td>
<td>-0.22</td>
</tr>
<tr>
<td>Cordes et al. (2010)</td>
<td>32</td>
<td>left DLPFC</td>
<td>10</td>
<td>110% MT</td>
<td>10000</td>
<td>10</td>
<td>0.30</td>
</tr>
<tr>
<td>Diabac-de Lange et al. (2013a,b)</td>
<td>32</td>
<td>bilateral DLPFC</td>
<td>10</td>
<td>90% MT</td>
<td>60000</td>
<td>15</td>
<td>0.25</td>
</tr>
<tr>
<td>Fitzgerald et al. (2008)</td>
<td>20</td>
<td>bilateral DLPFC</td>
<td>10</td>
<td>110% MT</td>
<td>30000</td>
<td>15</td>
<td>0.62</td>
</tr>
<tr>
<td>Goyal et al. (2007)</td>
<td>10</td>
<td>left DLPFC</td>
<td>10</td>
<td>110% MT</td>
<td>9600</td>
<td>10</td>
<td>2.22</td>
</tr>
<tr>
<td>Hajak et al. (2004)</td>
<td>20</td>
<td>left DLPFC</td>
<td>10</td>
<td>110% MT</td>
<td>10000</td>
<td>10</td>
<td>1.05</td>
</tr>
<tr>
<td>Holli et al. (2004)</td>
<td>22</td>
<td>left DLPFC</td>
<td>10</td>
<td>100% MT</td>
<td>10000</td>
<td>10</td>
<td>-0.47</td>
</tr>
<tr>
<td>Jin et al. (2012)</td>
<td>45</td>
<td>individual EEG alpha (8-13 Hz)</td>
<td>1</td>
<td></td>
<td></td>
<td>10</td>
<td>0.13</td>
</tr>
<tr>
<td>Klein et al. (1999)</td>
<td>31</td>
<td>right DLPFC</td>
<td>1</td>
<td>110% MT</td>
<td>1200</td>
<td>10</td>
<td>0.05</td>
</tr>
<tr>
<td>Li et al. (2016)</td>
<td>47</td>
<td>left DLPFC</td>
<td>10</td>
<td>110% MT</td>
<td>30000</td>
<td></td>
<td>0.23</td>
</tr>
<tr>
<td>Mogg et al. (2007)</td>
<td>17</td>
<td>left DLPFC</td>
<td>10</td>
<td>110% MT</td>
<td>20000</td>
<td>10</td>
<td>0.22</td>
</tr>
<tr>
<td>Novak et al. (2006)</td>
<td>16</td>
<td>left DLPFC</td>
<td>20</td>
<td>90% MT</td>
<td>20000</td>
<td>10</td>
<td>-0.29</td>
</tr>
<tr>
<td>Prikryl et al. (2007)</td>
<td>22</td>
<td>left DLPFC</td>
<td>10</td>
<td>110% MT</td>
<td>22500</td>
<td>15</td>
<td>1.13</td>
</tr>
<tr>
<td>Prikryl et al. (2013)</td>
<td>40</td>
<td>left DLPFC</td>
<td>10</td>
<td>110% MT</td>
<td>30000</td>
<td>15</td>
<td>1.33</td>
</tr>
<tr>
<td>Quan et al. (2015)</td>
<td>117</td>
<td>left DLPFC</td>
<td>10</td>
<td>80% MT</td>
<td>16000</td>
<td>20</td>
<td>0.40</td>
</tr>
<tr>
<td>Rabany et al. (2014)</td>
<td>30</td>
<td>mainly left DLPFC, also weaker stimulation right</td>
<td>20</td>
<td>120% MT</td>
<td>33600</td>
<td>20</td>
<td>0.30</td>
</tr>
</tbody>
</table>

* First author and year of publication.

b Placebo group, N = 15, 1 Hz group N = 17, 10 Hz group N = 16.

Fig. 3. Forest plot of effect sizes for active versus sham tDCS treatment of negative symptoms.
10 Hz rTMS were analysed separately, a mean weighted effect size of 0.43 (0.18–0.69) was observed (k = 12, total N = 557). Six studies (total N = 194) used rTMS protocols with more than 30,000 stimuli, their mean weighted effect size was 0.42 (0.00–0.84). When the analysis was limited to only studies that applied left prefrontal rTMS the mean weighted effect size was 0.36 (0.11–0.61; k = 13, total N = 569). When the analysis was limited to studies that stimulated above a motor threshold of 100%, the mean weighted effect size was 0.45 (0.20–0.69; k = 12, total N = 479). Meta-analysis of studies with a duration of treatment of longer than 2 weeks yielded a mean weighted effect size of 0.40 as compared to sham stimulation (0.16–0.64; k = 11, total N = 538).

We also compared studies on four other moderator variables of interest: age, duration of illness, number of rTMS stimuli (pulses) per week, and proportion of male patients included in the study. Whereas studies with younger patients than the mean of 39.1 years (k = 12, total N = 443) represented a mean effect size in the moderate range, 0.46 (0.14–0.78), studies with older patients than the mean (k = 9, total N = 335) had a small mean effect size of 0.26 (0.03–0.49). For studies with a mean shorter duration of illness of less than 13 years (k = 9, total N = 234) the mean effect size was 0.56 (0.21–0.92), while for studies with a longer duration of illness (k = 9, total N = 320) this was 0.29 (0.06–0.51). For studies that applied equal or more than 7500 stimuli per week the mean effect size was 0.41 (k = 9, total N = 249; 0.05–0.76), whereas for studies with less than 7500 stimuli per week this was 0.25 (k = 8, total N = 427; 0.03–0.47). Studies with more than 65% male participants reported a mean effect size of 0.41 (k = 13, total N = 475; 0.09–0.72). Studies with less than 65% male participants reported a mean effect size of 0.33 (k = 8, total N = 303; 0.11–0.56).

3.3. tDCS studies only

Information regarding the included studies applying tDCS is given in Table 3. Separate meta-analysis of tDCS studies showed a mean weighted effect size of 0.50 for actual stimulation versus sham (0.07–1.07; I² = 62%, k = 5, total N = 134), see Fig. 3. Due to the small amount of studies, no moderator analyses were possible.

3.4. rTMS and tDCS studies pooled together

The mean weighted effect size for all frontal noninvasive neurostimulation studies together (i.e., rTMS and tDCS studies combined in comparison to sham stimulation) was 0.61 (95% confidence interval 0.33–0.89; k = 27, total N = 961). The test for heterogeneity was significant (Q(26) = 109.3, p < 0.0001). Justifying the use of a random effects model, the I² statistic indicated that 77% of the heterogeneity between studies could not be accounted for by sampling variability. We also conducted an analysis without the outliers (Goyal et al., 2007 and Zhao et al., 2014). This analysis (k = 23, total N = 860) showed a mean weighted effect size of 0.35 (0.16–0.53). The I² statistic changed to 38%. Heterogeneity was significantly reduced and only marginally significant, Q(22) = 35.36, p = 0.04.

3.5. Sham stimulation

Analysis of sham rTMS (baseline - posttreatment comparison) showed a significant improvement of negative symptoms, d = 0.31 (0.09–0.52; I² = 0%, k = 16, total N = 333).

4. Discussion

This meta-analysis of 24 published studies (including 27 independent effect sizes) revealed a significant effect of non-invasive neurostimulation through rTMS or tDCS compared to sham stimulation (placebo). The magnitude of the effect size was in the moderate range. Separate analysis of rTMS and tDCS revealed moderate effect sizes for

---

### Table 3: Studies included in the meta-analysis applying rTMS and tDCS

<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>Rating scale</th>
<th>Location</th>
<th>Electrode types</th>
<th>Stimulation level</th>
<th>Duration, days</th>
<th>Sham protocol</th>
<th>Eft size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brunelin et al. (2012)</td>
<td>30</td>
<td>PANSS</td>
<td>a: F3/FP1 (l DLPFC); c: T3/P3 (temp-par)</td>
<td>35 cm² sponges 2 mA for 20 min 10 s, twice a day for 5 days</td>
<td>10 s, twice a day for 5 days</td>
<td>10 days</td>
<td>0.54 parallel</td>
<td>1.07 parallel</td>
</tr>
<tr>
<td>Fitzgerald et al. (2014)</td>
<td>24</td>
<td>SANS, PANSS</td>
<td>a: F3 or F3&amp;F4 (l/bilat DLPFC); c: TP3 or TP3&amp;TP4 (l/bilat temp-par)</td>
<td>35 cm² sponges 2 mA for 20 min 15 s, 15 days ramp up of real stimulation &amp; 30 s real stimulation</td>
<td>15 min continuous pulse</td>
<td>15 min</td>
<td>0.13 parallel</td>
<td>0.34 parallel</td>
</tr>
<tr>
<td>Mondlozh et al. (2016)</td>
<td>28</td>
<td>SANS</td>
<td>a: F3 (l DLPC); c: Fp2 (r orbitofr)</td>
<td>35 cm² sponges 2 mA for 20 min 10 s usage of dual-mode tDCS</td>
<td>10 s, twice a day for 5 days</td>
<td>40 sec onset real stimulation, every 550 ms over 40 sec</td>
<td>0.84 parallel</td>
<td>1.13 parallel</td>
</tr>
<tr>
<td>Smith et al. (2015)</td>
<td>37</td>
<td>SANS, PANSS</td>
<td>a: F3 (l DLPC); c: Fp2 (r orbitofr)</td>
<td>5.08 cm² sponges 2 mA for 20 min 5 s</td>
<td>5 s</td>
<td>40 sec onset real stimulation</td>
<td>0.56 parallel</td>
<td>0.54 parallel</td>
</tr>
</tbody>
</table>

Abbreviations: a = anodal; c = cathodal; temp-par (junct) = temporo-parietal (junction); r orbitofr = right orbitofrontal; s = sessions.
both, but this failed to reach statistical significance for the tDCS analysis, presumably because of the considerably smaller amount of studies and participants. Excluding outlier studies (with effect sizes > 2.0) from the rTMS meta-analysis, yielded a substantially smaller effect size (0.35) that was nonetheless significant. Thereby, heterogeneity was reduced significantly, indicating that the remaining studies were more consistent with each other regarding the estimation of effect magnitude. Exclusion of studies with unusually large effect sizes may represent an overly conservative approach, as they also belong to the peer-reviewed body of published evidence. However, it does imply that there are considerable differences between studies in terms of rTMS effects and that the overall effect can currently not be regarded to be stable and robust. This calls for an in-depth investigation of moderator variables that could contribute to such differences. Factors such as duration of treatment, variation in rTMS protocols, e.g. concerning intensity of stimulation (as expressed by percentage of the motor threshold) and patient characteristics could be relevant in this regard. On the other hand, false positive findings due to chance can also not be excluded as explanation for outliers, especially considering the relatively small number of participants in most studies.

Whereas previous meta-analyses were underpowered (cf. Table 1), our meta-analysis had a power of 0.87 to even establish a small effect. The results of our analyses clearly support the further development of noninvasive brain stimulation over the frontal cortex as a treatment for negative symptoms, as the mean weighted effect size remained significant even after removing studies with very large effect sizes. This may also imply that the observed effect size is robust against possible publication bias, as the remaining studies did not report large effects. Our additional analyses also suggest moderating factors that could be taken into account with regard to optimizing effects of brain stimulation. More specifically, for rTMS, high frequency stimulation with a protocol containing more than 7500 stimuli per week at an intensity of > 100% motor threshold, may be more effective than other protocols. The treatment may be more effective in younger patients with a shorter duration of illness, where the effect size was in the moderate range, in contrast to older-than-average patients, where a small effect size was reported. It could be suggested that there is more room for neuroplasticity in young people and people with a shorter duration if illness.

With regard to side of stimulation, it should be noted, that only one study (applying low-frequency stimulation) has investigated stimulation of the right DLPFC solely (Klein et al., 1999), thus this awaits further investigation. The efficacy of theta-burst stimulation also awaits firm conclusions, as there is not a sufficient amount of studies to warrant separate meta-analysis. Published studies typically did not report follow-ups after one month or more posttreatment. Thus, no conclusions can be drawn regarding duration of effects after the treatment, which is a notable limitation. Dlabac-de Lange et al. (2015a) reported a stable reduction of negative symptoms that was still present 3 months post-treatment. Future studies should by default include follow-up measurements.

A separate analysis of sham conditions (pre- versus posttreatment) yielded a significant effect size of 0.31. It should be noted that this is not comparable to the effect sizes obtained for verum stimulation, as those were over and above sham effects. Nonetheless, it indicated that a placebo-effect occurs, as is common in medical and psychological treatments. Indeed, a recent meta-analysis of sham conditions in rTMS studies of auditory hallucinations in schizophrenia (Dollfus et al., 2016) also observed a significant effect size of 0.29 (21 studies), which is almost identical to the effect size we observed. The lack of heterogeneity in our analysis of sham effects indicates a high consistency across studies of this effect. Most studies used a sham condition in which the coil was rotated (with 45 or 90 degrees) away from the scalp, such that the side of the coil maintained contact with the scalp but the magnetic field was directed away from the brain. Even though many patients can not easily distinguish this condition from real stimulation, it is not an ideal sham condition. That is, verum stimulation induces scalp sensations that are not (or almost not) present in these sham conditions. Currently, sham coils are available with a cutaneous electrical stimulator that mimics the sensation on the scalp. Together with a parallel group design (in which patients don’t get both real and sham stimulation which allows them to compare differences), we would advocate use of such sham coils.

It should be noted that further possible benefits of frontal neurostimulation have been recently highlighted, specifically with regard to cognitive functioning (for review see Enriquez-Geppert et al., 2013). Thus, prefrontal neurostimulation may also improve other aspects of information processing abnormalities in schizophrenia. Indeed, a preliminary finding of an improvement in verbal fluency performance after rTMS over the DLPFC (bilaterally) was reported by Dlabac-de Lange et al. (2015a). Verbal fluency is thought to depend in part on executive functioning subserved by prefrontal circuits (Roehrich-Gascon et al., 2015). In addition, a recent study suggested that rTMS over the left DLPFC may reduce EEG-measured hypofrontality (Kamp et al., 2016). An fMRI study of activation during a planning task reported increased frontal activation after bilateral DLPFC stimulation with rTMS in schizophrenia patients (Dlabac-de Lange et al., 2015b). It should be noted, though, that the number of patients that could be included in this study, was relatively low (24 patients divided over two groups), underlining the need for replication.

Some methodological issues deserve discussion. First, measurement of negative symptoms was generally accomplished with the use of the SANS or the negative subscale of the PANSS. It should be noted that the SANS is more comprehensive and has been shown to be sensitive to change in pharmacological trials (Strous et al., 2003). In addition, in recent years measures have been developed that also assess experiential aspects of negative symptoms, e.g. CAINS (Kring et al., 2013) and BNNS (Kirkpatrick et al., 2011). No brain stimulation trials using these measures have been reported as yet. Another methodological issue regards whether the study concerns a monocenter trial or a multicenter trial. A clear advantage of multicenter trials is the potential for including a larger sample, as was the case for the study by Wobrock et al. (2015), which is the largest rTMS trial of negative symptoms to date, involving three centers. An advantage of monocenter trials, however, may be that it is more feasible to keep execution of procedures identical, as patients may be seen by the same researchers who communicate more among each other on a daily basis. The need of studies with larger samples is so compelling however, that multicenter trials are to be preferred, whilst ensuring standardization of procedures across sites. A final methodological issue concerns the heterogeneity of findings across studies. Heterogeneity is a hallmark of psychotic disorders and partly an artefact of diagnostic systems that allow for considerable differences in psychopathological presentation within one category. In addition to such symptomatic heterogeneity (e.g. some patients have hallucinations in addition to negative symptoms, others only delusions, others both), there is heterogeneity in comorbidities, severity of illness, duration of illness, type of treatment etc. It would be of interest to conduct studies in selected populations, such as first-episode patients. They can already present with negative symptoms and treatment may prevent further deterioration.

In conclusion, the results of our meta-analysis show that noninvasive neurostimulation can improve negative symptoms in patients with schizophrenia. For the analysis on rTMS trials, even after excluding two studies with extreme effect sizes, a significant mean effect size of 0.31 remained (based on 18 studies) and heterogeneity was nonsignificant, indicating consistency across studies. Our analyses furthermore suggested that protocols with high frequency stimulation containing more than 7500 stimuli per week at an intensity of > 100% motor threshold, may be more effective than other protocols. The treatment may be more effective in younger patients with a shorter duration of illness. However, protocols with frequencies other than 10 Hz and locations other than the left DLPFC have been studied less
References


Enriquez-Geppert, S., Huster, R.J., Herrmann, C.S., 2013. Expressive deficits and social-emotional withdrawal - to a different degree. Novel measures of negative symptoms may also be included as outcome measures, as they may be more comprehensive (e.g., Krings et al., 2013). Future studies should also investigate the neural basis of noninvasive neurostimulation treatments in more detail, which may yield insights into its underlying mechanisms and clues for more targeted interventions.

Acknowledgment

AA was supported by a VICI grant from the Netherlands Organisation for Scientific Research (N.W.O.) grant no. 453.11.004 and an ERC consolidator grant (project no. 312787).


