DEVELOPING A PROSPECTIVE EFFICACY ASSESSMENT FOR EEG NEUROFEEDBACK-BASED LANGUAGE THERAPY
PRESENTATION OVERVIEW

Context
- Language therapy & neurofeedback
- Inefficacy problem
- Neurofeedback protocol development

Current Work
- Dyslexia
- Aphasia
- Prospective NF Efficacy Assessment

Future Directions
TREATMENT OF LANGUAGE DISORDERS

Primarily handled by speech-language pathologists
- Sometimes there are interdisciplinary teams, as in acute brain injury

Focus is on behavioral, speech-language interventions
- Find the speech-language deficit, work on those skills
- Cognition is considered, but generally only as a secondary concern

Treatment is rarely informed by individual brain assessments (MRI, qEEG, etc.)
- There are few ways to tie individual brain assessment to treatment planning
- Average brain patterns are often limited in their practical application to individuals – individual variability is high, especially in brain injury
The crucial role of brain networks is a major focus of current brain-language research, and neurofeedback is one of the only methods for treating network dysfunction.

Multiple populations with co-morbid language deficits have been treated with neurofeedback:
- TBI, Stroke, Dyslexia, Autism, etc.

Though there have been several published examples of secondary language improvement with neurofeedback (and many anecdotal reports), it is generally not the primary target:
- Koberda, 2015 (TBI); Coben & Padolsky, 2007 (Autism).
NEUROFEEDBACK-RELATED IMPROVEMENTS

Assumption: NF -> Targeted brain changes -> Behavioral improvement

Whether practitioners gauge outcomes by self-report measures of symptom improvement, or by objective testing, there are several possible causes of improvement in behavior:

- 1: Targeted brain changes -> behavioral improvement
- 2: Non-targeted brain changes -> behavioral improvement
- 3: Non-specific effects (e.g. placebo) -> behavioral improvement

How do you feel now?
Several recent high-profile papers have questioned NF’s causal role (Thibault & Raz, 2018, Schabus et al., 2017 & associated responses)

- **Specificity problem:** In most contexts, there are not easy methods to assess correlated brain changes (e.g. coherence shifting with power training, or vice versa; Walker & Horvat, 2010) – c.f. causality of targeted vs non-targeted brain changes

- **Methods problems:** NF methods continue to evolve rapidly, which is fantastic, but often leaves little time for rigorous testing, and establishment of the evidence-base (c.f. Coben, Hammond & Arns, 2018)

- **Inefficacy problem:** In many contexts, patients aren’t able to successfully modulate brain activity – estimates range from 15 – 60 % (Alkoby et al, 2018) suggesting that this “responder/non-responder” dynamic is likely a larger problem than we realize.
REASONS FOR INEFFECTIVENESS?

- **Amount of Control**
  - **Traits**
    - Individual differences - brain damage, normal structure/function variance, genetic factors (e.g. BDNF genotype)
  - **States**
    - Alertness, attention, medications, etc.
  - **Methods**
    - Feedback method (e.g. Raw/Z-Scores, LORETA), sites/bands trained, signal processing algorithms, etc.
**LANGUAGE-SPECIFIC PROTOCOL DEVELOPMENT**

Method 1: Targeting neural patterns common to a particular *disorder*

- Scientific literature can provide good information on relevant brain areas/patterns
- However, individuals may not follow “average” patterns
  - Disorders are not homogenous – individual variability in symptoms matters
  - Protocols can be individually “matched”, via symptom-checklist approaches
- Also, many areas/patterns are defined based on MRI and other methods
  - EEG specificity matters – frequency & metric information is important, as is location

---

**Method**: Average EEG measures across the disordered group – here, the group is 50 children with dyslexia.

- Highlights global high beta activity.
- Q: Should we downtrain beta?

*(Fillmore & Ritter, 2017)*
Method 2: Target neural patterns common to specific symptoms

- Less well-defined in the literature, but has potential to be more targeted
- Still subject to concerns about EEG-specificity and individual variability (c.f. symptom-checklists)

<table>
<thead>
<tr>
<th>Delta</th>
<th>Theta</th>
<th>Alpha</th>
<th>Beta</th>
<th>High Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Nonword Spelling

Method: Correlate EEG measures with relevant behavioral measures

Correlations seen between higher global power in low-mid (theta, alpha) frequencies and higher performance for both non-word spelling and written narratives

Q: Should we uptrain global alpha/theta?
Method 3: Target neural patterns which change in relation to specific symptom improvement

- Even less well-defined in the literature, but has potential to be very precisely targeted
- Requires significant work to develop evidence-based protocols

**Nonword Spelling**

<table>
<thead>
<tr>
<th>Delta</th>
<th>Theta</th>
<th>Alpha</th>
<th>Beta</th>
<th>High Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Delta" /></td>
<td><img src="image2" alt="Theta" /></td>
<td><img src="image3" alt="Alpha" /></td>
<td><img src="image4" alt="Beta" /></td>
<td><img src="image5" alt="High Beta" /></td>
</tr>
</tbody>
</table>

**Written Narrative Retell - # T-units**

<table>
<thead>
<tr>
<th>Delta</th>
<th>Theta</th>
<th>Alpha</th>
<th>Beta</th>
<th>High Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Delta" /></td>
<td><img src="image2" alt="Theta" /></td>
<td><img src="image3" alt="Alpha" /></td>
<td><img src="image4" alt="Beta" /></td>
<td><img src="image5" alt="High Beta" /></td>
</tr>
</tbody>
</table>

**Method:** Correlate EEG measures with improvement in relevant behavioral measures

Correlation between high frontal beta and better performance, and also posterior alpha for written narratives. Note that now the difference from normal (increased high beta) in the group as a whole now seems it may adaptive.

Q: Should we uptrain frontal beta?
APHASIA CASE STUDY

67 yr. old male, left CVA resulted in coma, progressed to anomic aphasia by 2 yrs post – GM/WM damage

24* sessions of Z-score Surface NF – DSI-24 System

- Each 20 min NF session was done in conjunction with speech therapy for picture naming and perseveration
- *Began with 4 sessions of sham feedback to establish a baseline, recorded EO rest at each session
- Looked at LPR results for LORETA Tx, but they were hard to interpret, given brain damage
- Targets: T3 & T4 Theta power (All-or-none Z-thresh=±1.5)
There was significant improvement for perseveration, but not correct score or total score.

Most of the improvement was in the first few sessions before real NF began – if you look at improvement during the NF period, none of the measures improved significantly.
No metrics showed significant learning over time, either for QEEG or NFB
STUDY OUTCOMES

- Some behavioral improvement, but it was not clearly tied to NF
- No clear evidence of EEG learning in NF over time in trained metrics
- Some changes in the QEEG from pre to post were seen (mainly in beta), but largely outside of the trained metrics
- Why not a better outcome?
  - Not enough training?
  - Wrong training methods/metrics?
  - Was patient a “non-responder”?
HOW TO PREDICT TREATMENT RESPONSE?
DESIGNING A PROSPECTIVE EFFICACY ASSESSMENT

Key questions
- How can we tell who will respond well to treatment?
  - Many possible predictors – Brain structure/function (MRI), QEEG, Cognitive function, etc.
  - Likely depends largely on specific treatment goals/methods
- How do we know how to tailor treatments to be most effective?
  - Are all sites & metrics equally trainable? Are all NF approaches equally effective?

Study design
- Two phases: 1) Surface NF, 2) LORETA NF - QEEG done pre-post for each
- Pre-post cognitive testing using the NIH Toolbox
- Initial comparison of NF for several common metrics, at sites within the language network – done both for surface and LORETA methods
- Follow-up sessions – more in-depth study of particular sites/metrics
- 1 hour sessions, 4x/week
- Nine young, healthy participants (continued data collection ongoing)
INITIAL ASSESSMENT

- Z-score training in Neuroguide – All-or-none Z-thresh=±2.0
- 3 min blocks for Surface, 5 min blocks for LORETA
- Simple auditory & visual feedback
- Eyes-open rest done at beginning and end of sessions

### Surface Neurofeedback

<table>
<thead>
<tr>
<th>Block</th>
<th>Site</th>
<th>Band</th>
<th>Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>All</td>
<td>All</td>
<td>Resting Eyes Open - 5 min.</td>
</tr>
<tr>
<td>1</td>
<td>T3</td>
<td>Theta</td>
<td>Absolute Power</td>
</tr>
<tr>
<td>2</td>
<td>T3-T4</td>
<td>Theta</td>
<td>Interhemi AmpAsym</td>
</tr>
<tr>
<td>3</td>
<td>T3-F7</td>
<td>Theta</td>
<td>Intrahemi AmpAsym</td>
</tr>
<tr>
<td>4</td>
<td>T3-T4</td>
<td>Theta</td>
<td>Interhemi Coherence</td>
</tr>
<tr>
<td>5</td>
<td>T3-F7</td>
<td>Theta</td>
<td>Intrahemi Coherence</td>
</tr>
<tr>
<td>6</td>
<td>T3-T4</td>
<td>Theta</td>
<td>Interhemi Phase Lag</td>
</tr>
<tr>
<td>7</td>
<td>T3-F7</td>
<td>Theta</td>
<td>Intrahemi Phase Lag</td>
</tr>
<tr>
<td>B</td>
<td>All</td>
<td>All</td>
<td>Resting Eyes Open - 5 min.</td>
</tr>
</tbody>
</table>

### LORETA Neurofeedback

<table>
<thead>
<tr>
<th>Block</th>
<th>Site</th>
<th>Band</th>
<th>Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>All</td>
<td>All</td>
<td>Resting Eyes Open - 5 min.</td>
</tr>
<tr>
<td>1</td>
<td>L_BA22</td>
<td>Theta</td>
<td>Absolute Power</td>
</tr>
<tr>
<td>2</td>
<td>L_BA22 - R_BA22</td>
<td>Theta</td>
<td>Interhemi Coherence</td>
</tr>
<tr>
<td>3</td>
<td>L_BA22 - L_BA44</td>
<td>Theta</td>
<td>Intrahemi Coherence</td>
</tr>
<tr>
<td>4</td>
<td>L_BA22 - R_BA22</td>
<td>Theta</td>
<td>Interhemi Phase Lag</td>
</tr>
<tr>
<td>5</td>
<td>L_BA22 - L_BA44</td>
<td>Theta</td>
<td>Intrahemi Phase Lag</td>
</tr>
<tr>
<td>B</td>
<td>All</td>
<td>All</td>
<td>Resting Eyes Open - 5 min.</td>
</tr>
</tbody>
</table>
Based on initial QEEG’s, most metrics were within normal range, with the exception of intrahemispheric coherence.

Based on reward proportions from NF blocks, there were large differences in difficulty across metrics – this was much more than expected, especially for intrahemispheric phase lag.
STUDY RESULTS — LORETA NF

Based on initial QEEG’s, all metrics were within normal range.

Based on reward proportions from NF blocks, there were large differences in difficulty across metrics – absolute power was the most difficult, and intrahemispheric phase lag (hardest for surface) was at ceiling.
WITHIN-SESSION LEARNING – SURFACE VS LORETA

Surface - No significant learning
• This was similar for one metric at a time and multiple metrics at a time
• Training multiple surface metrics was much more difficult than a single metric

LORETA - Good learning
• This was true for both one metric at a time and multiple metrics at a time
• Training multiple surface metrics was similar to training a single metric
SUMMARY OF CURRENT WORK

Developing language-specific protocols to link brain improvements to symptom improvements is key, but much more work is needed.

We need a better understanding of the inefficacy problem, and how it relates to difficult populations, such as aphasia.

In comparing NF for different metrics and methods, we find some unexpected results, which need to be considered in how we develop NF protocols.