Neuroscience of Deception: Limitations of fMRI

Craig Stark, Ph.D.
University of California, Irvine

Traditional Polygraph

Q: Should polygraph data be admissible?

- Implied accuracy and authority
- Technological revolution. Machines can do anything for us
- Machines don’t lie. They are cold and impartial.
- “Reading” a polygraph is an art, not a science. Designing the questions is too.

Q: Is a brain scanner any different?

The New York Times

“In response to images of Democratic candidates, men exhibited activity in the medial orbital prefrontal cortex, indicating emotional connection and positive feelings.”

“Images of Fred Thompson led to increased activity in the inferior frontal cortex, a brain structure associated with empathy.”

“Subjects who had an unfavorable view of John Edwards responded to pictures of him with feelings of disgust, evidenced by increased activity in the insula, a brain area associated with negative emotions.”
“We can read your mind”

From J. Illes, UBC

http://www.wired.com/wiredscience/2009/03/noliemri/


“Lure of lie detectors spooks ethicists”

“Stanford Center for Law & the Biosciences Blog”
Overview

- Basics of fMRI
- General issues and challenges associated with using fMRI for studying the mind
- Challenges facing fMRI lie detection
- Sample studies
- Can the test be beaten? Do the tests generalize?

Issues

- fMRI is very noisy
  - Typically many trials and many subjects
- fMRI is contrastive. There is no baseline.
  - Can only see X vs. Y. Effects can come from X or Y and you won’t know which. Both can be active and you won’t see either.
- Need an appropriate contrast and control.
- Together:
  - Null results are trivial to obtain as are artifactual results
  - Very difficult to answer “Big Questions”

fMRI: Basics

**Brain**

- Increased brain activity produces a local increase in the blood oxygenation

**Physics**

- Oxygenated and deoxygenated blood have different magnetic properties

MR images are brighter (there is more signal) when there is greater activity

The BOLD Effect

- Buxton (2001)

Menon et al. (1995)
What (arbitrary) threshold should we use?
Groups vs. Individuals

Issues
- fMRI is very noisy
  - Typically many trials and many subjects

- Together:
  - Null results are trivial to obtain as are artifactual results
  - Very difficult to answer “Big Questions”

Rest / Baseline Assumption
- Because no task is being performed during rest, rest represents something akin to a “zero-activity”, baseline condition.
- If a region’s activity is the same in rest as in a task, the region is not involved in the task.
- Sadly, the assumption is often invalid.

*Failure to control thought processes (subjects doing what they want to not what they are told to do) distorts or obliterates the data.*

Non-Mnemonic Task Activity

- Novel Pictures
- Familiar Pictures
- Moving Fixation
- Arrows
- Noise Detection
- Odd/Even Digits
- Rest
MTL Activity vs. Baseline Task

Issues
- fMRI is very noisy
  - Typically many trials and many subjects
  - fMRI is contrastive. There is no baseline.
  - Can only see X vs. Y. Effects can come from X or Y and you won’t know which. Both can be active and you won’t see either.
  - Need an appropriate contrast and control.

Challenge: Individual Differences at the Neural level
- Individual functional and structural anatomy shows significant differences even when demographics are well-controlled.
- Significant environmental effects on the brain
- Significant disease / drug use / mental health effects on the brain and fMRI signal.
- Each of these are still largely unknown and how to counteract for them is entirely unknown.

We cannot use “normals” as a basis for comparison. Activity differences may be due to ancillary differences.

Challenge: Individual Differences at the Psychological Level
- Individual differences in the way people react to being falsely accused.
- fMRI wants ~30 of the same trial type to have any kind of reliable estimate.
- Individual differences in adaptations (changes in the reactions) across many trials.
- Even if adaptation were similar across people, each “Did you do X” may well be different.

These differences lead to increased variance both within and between subjects. Variance lowers SNR leading to failure to observe activity or differences.
Challenge: Ethological Validity

- Typical tasks in controlled tests lack ethological validity
- Lie about card you’re holding to win $20, lie about unremarkable made-up autobiographical event, etc.
- Little if any emotion
- Emotional reactions have large effects on the brain that can be long-lasting and variable across people.
- May be a poor basis for determining whether you’re lying about murdering your child.
- Untested in those skilled at deception

We do not know how this will work in real life situations.

Challenge: Detection Sensitivity and Base Rates

<table>
<thead>
<tr>
<th></th>
<th>$p(Liar) = 0.5$</th>
<th>$p(Liar) = 0.001$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sensitivity/</td>
<td>Specificity</td>
</tr>
<tr>
<td></td>
<td>$p(\text{Lie}</td>
<td>\text{Truth})$</td>
</tr>
<tr>
<td>Cards</td>
<td>~0.72</td>
<td>0.13</td>
</tr>
<tr>
<td>Personal Items</td>
<td>~0.8</td>
<td>0.1</td>
</tr>
<tr>
<td>Mock Crime</td>
<td>~0.85</td>
<td>0.08</td>
</tr>
</tbody>
</table>

$Wolpe$ et al. (2005), adapted from Ben-Shakhar & Elaad (2003)

$p$ of condition affects rates of real true and false positives

Literature

- ~18 peer-reviewed publications since 2001
- Instruct or give participants incentive to lie on some trials and not others (e.g., win $20 if you can conceal holding a card).
- College undergrads, not criminals or experts at deception.
- Contrast Lie trials with Truth trials to identify brain regions that show differences.
- Frontal (executive / higher-level cognition) and cingulate (error / conflict) often found here (and in many other studies).
- Pattern Classification: New, advanced approach to data analysis that is well-suited to this problem.

Pattern Classification

- Train a computer model on a portion of the data
- Tell it “lie” and “truth” trials
- It learns what lies and truth “look like” in a complex, non-linear, high-dimensional way.
- No need to identify specific locus of lie vs. truth
- Give it new data (unused trials from the same subject or data from a new subject) and have it classify the trials as “lie” or “truth”
Sample Studies

- Given 2 cards and $20 outside the scanner. Told to deny having 1 of the cards when asked in the scanner.
- Both of those and other cards shown.
- Do you have this card?
- Told they could keep the $20 only if they could successfully lie about the one card.
- Consistently said “yes” to the Truth card and “no” to the Lie card.
- Two studies from one lab

Results

- Study 1: Davatzikos et al. (2005) NeuroImage
  - Train on 99% of the data, test on 1%
  - 99% accurate within-subject at predicting the 1% trials left out from the analysis
  - 88% accurate at predicting a “new” subject
- Study 2: Langleben et al. (2005) Human Brain Mapping
  - 78% accurate at predicting single trials

Predictive Brain Regions

<table>
<thead>
<tr>
<th>Study 1</th>
<th>Study 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right inferior prefrontal cortex (inf. and sup. frontal gyri BA 6 and 44)</td>
<td>Left inferior parietal lobule (BA 40)</td>
</tr>
<tr>
<td>Bilateral superior temporal and inferior parietal gyri (BA 38 and 40)</td>
<td>Left middle frontal gyrus (BA 6)</td>
</tr>
<tr>
<td>Bilateral pericentral area (BA 2, 3 &amp; 4)</td>
<td>Left inferior frontal gyrus (BA 45)</td>
</tr>
<tr>
<td></td>
<td>Right cuneus (BA 18)</td>
</tr>
<tr>
<td></td>
<td>Right middle temporal gyrus (BA 37)</td>
</tr>
<tr>
<td></td>
<td>Left cerebellum</td>
</tr>
</tbody>
</table>

Conclusions

- High degree of accuracy possible at learning to classify trials as lies or truth.
- But...
  - Compliant subjects
  - Homogeneous subject pool
  - Tasks well-controlled (scientists will like it), but ethological validity is fairly low
  - Predictive regions across very similar tasks with similar methods from the same lab showed only partial overlap across studies.
  - With other methods, same region has been labeled Truth > Lie and Lie > Truth across studies
Beating the Test

- Exploit the low SNR and need for consistency across trials.
- Small physical motions introduce variance
- Do them during “control lie” trials but not “control true” or “critical lie”.
- Or, just do them and all the data goes to noise
- Have a brain that is abnormal

Beating the Test

- Mentally introduce variance
  - Think of other things during the trials (treat like that “Rest” task)
  - Even if you must think of the critical information first, think of something else immediately and exploit the low temporal resolution (e.g., something novel).
  - Think of “lies” (even the critical one) during a number of the Truth trials.

Myth Busters #93: Beat the MRI

- Steal a watch or a ring
- Beat Cephos’ MRI, get $1000. Loose and take a 3,000 mile bus trip back to San Francisco

Tori Kari Grant
Think happy / fearful thoughts Keep the brain always active

Myth: PLAUSIBLE

- Two were “caught”, one beat the test
- It called “ring” each time
- Odds of getting 2 or more heads in 3 coin tosses: 50% (Note 50/50 chance of lie vs. truth).
- One strategy worked 100% of the time and the other 0%.
- In a similar scenario, all were easily caught with a traditional polygraph.
Beating the Test

- Practice the critical lie
- True vs. false memories differ in their richness.
- Strong vs. weak memories differ in how automatic they are and how much is retrieved.
- Internally-generated memory encoding and externally-generated memory encoding are remarkably similar.
- Practice the lie enough and it will look a lot more like a “control truth” trial.

Conclusions

- fMRI is noisy and not well-suited to many questions.
- Differences in brain anatomy and functional anatomy make comparing across demographic groups virtually impossible.
- Existing studies are promising but have clear weaknesses.
- Nobody has investigated serious attempts to beat the test and there are apparent ways to do so (and some evidence these can be beaten).
- fMRI is not up to use as legal evidence, especially given the implied accuracy of high-tech brain scans.