Group Actively Open-Minded Thinking

Sydney Scott, Nick Rohrbaugh, Yuan Hou, Emlen Metz, Jon Baron, Philip Tetlock
Geopolitical Forecasting Tournament

- 50 teams, on average 14 active members each

- Probability judgments:
  - “Will China’s official annual GDP growth rate be less than 7.5 percent in Q1 2014?”
  - “How many additional countries will report cases of the Ebola virus as of 9 May 2014?”
  - “Between 2 April 2014 and 10 May 2014, will Russia officially annex any additional Ukrainian territory?”

- Probability judgments are evaluated using Brier scores

- Post comments in discussion forum

- Answer one survey about individual differences, three surveys about team characteristics
Actively Open-Minded Thinking (AOT)

1. Extensive, unbiased search for evidence and information

2. Unbiased inference from that information

People should take into consideration evidence that goes against their beliefs.

People should search actively for reasons why their beliefs might be wrong.
Open-Minded Thinking

$r = -0.11$, $p = 0.02$
Groups of Open-Minded Individuals

- Advantageous:
  - Groups can find **more information** by leveraging each member’s ability to search
  - Groups can make better use of that information by **unbiased inference**, focusing not just on shared info but the private info of each member
Groups of Open-Minded Individuals

\[ r = -0.04 \]
\[ p = 0.73 \]
Group Actively Open-Minded Thinking (GAOT)

Three rounds of data collection, in November 2013, March 2014 and June 2014.

Team members were asked to rate their team on items such as these:

1. We carefully discuss evidence that could cause us to change our minds.

2. When people express opinions that differ from the rest of the group, we ignore them. (reversed)

Round 1: 4 items, alpha = .72

Round 2: 16 items, alpha = .66

Round 3: 16 items, alpha = .69

Pairwise correlations between scores = .58 (1 & 2), .61 (1 & 3), and .71 (2 & 3)
Open-Minded Groups

$r = .12$

$p = .42$
Open-Minded Groups

$r = -.50$, $p = .0002$
Open-Minded Groups

- Multilevel Random Coefficient Model:

\[
\text{Brier Score}_{ij} = \beta_{0j} + \beta_{1j} (\text{Individual Open-Mindedness}_{ij}) + r_{ij} \\
\beta_{0j} = \gamma_{00} + \gamma_{01} (\text{Group Open-Mindedness}_{j}) + u_{0j} \\
\beta_{1j} = \gamma_{10} + u_{1j}
\]

- Significant fixed effects
  - Individual open-mindedness (\(\beta = -0.09, \ p = 0.01\))
  - Group open-mindedness (\(\beta = -0.50, \ p < 0.001\))
Open-Mindedness versus Past Performance

- **November 2013**
  - Individual open-mindedness ($\beta = -0.06$, $p = 0.14$)
  - Past performance ($\beta = 0.12$, $p < 0.001$)
  - Group open-mindedness ($\beta = -0.13$, $p = 0.02$)

- **March 2014**
  - Individual open-mindedness ($\beta = -0.11$, $p = 0.01$)
  - Past performance ($\beta = 0.24$, $p < 0.001$)
  - Group open-mindedness ($\beta = -0.31$, $p = 0.01$)
Open-Mindedness Individual Difference Correlates

- Individual Open-Mindedness
  - Males ($M_{\text{male}} = 5.90, M_{\text{female}} = 5.63, p < 0.001, d = 0.43$)
  - Higher IQ (Correlation Raven’s Matrices = .16, $p < 0.001$)
  - Higher Shipley’s Abstraction scores ($r = .13, p = 0.001$)
  - Higher CRT ($r = .22, p < 0.001$)
  - Higher numeracy scores ($r = .21, p < 0.001$)
  - Less hierarchical cultural world views ($r = -.20, p < 0.001$)

- Group Open-Mindedness
  - No average individual differences
  - No differences in variance
Open-Minded Discourse

- Individual Open-Mindedness
  - More comments ($r = .15, p < 0.001$)

- Group Open-Mindedness
  - More comments ($r = .48, p < 0.001$)
  - More replies to team members.
  - More comments with questions receive replies from team members.
  - Marginally more likely to ask questions and to post longer comments.
What we know:

- Individual AOT correlates with an individual’s accuracy.
- Group AOT correlates with a group’s accuracy.
- Group AOT and Individual AOT are separate constructs:
  - Group AOT and accuracy are not predicted by the average individual AOT of the members.
  - Group AOT and individual AOT correlate with different variables (roughly, team process and intelligence metrics respectively)
Future Directions

- Which processes are driving the effect of group AOT on forecasting accuracy? Search? Inference?

- How can we increase those processes via instilling actively open-minded evidentiary norms?
end
How do we create open-minded groups?

Lab experiments can help us discover how to maximize group open-mindedness.
Manipulating Group Open-Mindedness

- Hidden Profiles: Find the murderer.

  Detective A knows:
  a. 3 clues implicating Suzie
  b. 2 clues implicating Lisa
  c. 1 clue implicating John
d. 1 clue implicating John

  Detective B knows:
a. 3 clues implicating Suzie
b. 2 clues implicating Lisa
c. 1 clue implicating John
d. 1 clue implicating John

  Detective C knows:
a. 3 clues implicating Suzie
b. 2 clues implicating Lisa
c. 1 clue implicating John
d. 1 clue implicating John

  Pooled Shared Information indicates Suzie is the killer.

  Pooled Shared and Unique Information indicates John is the killer.
Manipulating Group Open-Mindedness

- Can we manipulate groups to be open-minded and search for evidence in an unbiased manner?

- Encourage in instructions “allowing every person’s ideas to be heard, including reasons why the group might be wrong” should increase extent of search, search for unique information and unbiased inference when unique information is discovered.
## Process Correlates (Group Level)

<table>
<thead>
<tr>
<th></th>
<th>GAOT (N = 50)</th>
<th>IAOT (N = 50)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IQ</td>
<td>-0.36</td>
<td>0.35</td>
</tr>
<tr>
<td>CRT</td>
<td>-0.09</td>
<td>0.45</td>
</tr>
<tr>
<td>Words</td>
<td>0.64</td>
<td>-0.16</td>
</tr>
<tr>
<td>Questions</td>
<td>0.68</td>
<td>-0.2</td>
</tr>
<tr>
<td>Links</td>
<td>0.63</td>
<td>-0.26</td>
</tr>
<tr>
<td>Prosociality</td>
<td>0.28</td>
<td>0.08</td>
</tr>
<tr>
<td>$ to self</td>
<td>-0.3</td>
<td>-0.02</td>
</tr>
<tr>
<td>$ to other</td>
<td>0.27</td>
<td>0.08</td>
</tr>
</tbody>
</table>
## Process Correlates Round 2 (Group Level)

<table>
<thead>
<tr>
<th></th>
<th>GAOT (N = 50)</th>
<th>IAOT (N = 50)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IQ</td>
<td>-0.22</td>
<td>0.35</td>
</tr>
<tr>
<td>CRT</td>
<td>0.13</td>
<td>0.45</td>
</tr>
<tr>
<td>Words</td>
<td>0.57</td>
<td>-0.16</td>
</tr>
<tr>
<td>Questions</td>
<td>0.57</td>
<td>-0.2</td>
</tr>
<tr>
<td>Links</td>
<td>0.59</td>
<td>-0.26</td>
</tr>
<tr>
<td>Prosociality</td>
<td>0.06</td>
<td>0.08</td>
</tr>
<tr>
<td>$ to self</td>
<td>-0.1</td>
<td>-0.02</td>
</tr>
<tr>
<td>$ to other</td>
<td>0.17</td>
<td>0.08</td>
</tr>
<tr>
<td>% Females</td>
<td>0.24</td>
<td>-0.08</td>
</tr>
</tbody>
</table>
## Process Correlates (Individual Level)

<table>
<thead>
<tr>
<th></th>
<th>GAOT (N = 190)</th>
<th>IAOT (N = 636)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IQ</td>
<td>-0.06</td>
<td>0.18</td>
</tr>
<tr>
<td>CRT</td>
<td>-0.01</td>
<td>0.33</td>
</tr>
<tr>
<td>Max Time</td>
<td>0.04</td>
<td>0.01</td>
</tr>
<tr>
<td>Words</td>
<td>0.28</td>
<td>0.03</td>
</tr>
<tr>
<td>Questions</td>
<td>0.31</td>
<td>0.02</td>
</tr>
<tr>
<td>Links</td>
<td>0.30</td>
<td>0.05</td>
</tr>
<tr>
<td>Questions Replied</td>
<td>0.12</td>
<td>0.10</td>
</tr>
<tr>
<td>Prosociality</td>
<td>0.02</td>
<td>-0.07</td>
</tr>
<tr>
<td>$ to self</td>
<td>0.01</td>
<td>0.12</td>
</tr>
<tr>
<td>$ to other</td>
<td>0.04</td>
<td>-0.05</td>
</tr>
</tbody>
</table>
Pilot 1

- 17 groups of 3 participants each interact in person

- First, as isolated individuals, then together in groups around one computer, participants complete:
  - CRT
  - Raven’s matrices
  - Forecasting task (0 1 0 0 1 0 1 0...)

- Then, individuals return to individual stations, report individual and group actively open-minded thinking
Group Actively Open-Minded Thinking

1. We follow our intuitions.
2. We seek evidence that contradicts our beliefs.
3. We discuss evidence indicating our beliefs are incorrect.
4. We change our minds in the face of new evidence.
5. We revise our beliefs in response to new information.
6. We persevere when our beliefs are challenged.
7. We disregard evidence that conflicts with our beliefs.
8. We are decisive as a group.
9. We achieve consensus quickly.
10. We prefer members to stay the course, not rock the boat.
11. We encourage each other to discuss different alternatives.
12. We encourage each other to dissent.
13. We do not tolerate much difference in opinion among our members.
14. We try to think through all the evidence and alternatives before making any decisions.
**Study 1 (N = 17)**

<table>
<thead>
<tr>
<th></th>
<th>Team IQ</th>
<th>Ave. Member IQ</th>
<th>Group AOT</th>
<th>Ave. Member IAOT</th>
<th>CRT Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ave. Member IQ</td>
<td>0.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group AOT</td>
<td>0.50*</td>
<td>0.16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ave. Member IAOT</td>
<td>0.40</td>
<td>-0.44†</td>
<td>0.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRT Scores</td>
<td>0.45†</td>
<td>0.36</td>
<td>0.47†</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>Brier Scores</td>
<td>-0.71**</td>
<td>-0.20</td>
<td>-0.43†</td>
<td>-0.51*</td>
<td>-0.29</td>
</tr>
</tbody>
</table>
Study 2

- 37 groups of 3 participants each interact in person

- First, as isolated individuals, then together in groups around one computer, participants complete:
  - CRT
  - Raven’s matrices

- Also reported individual actively open-minded thinking
## Study 2 (N = 37)

<table>
<thead>
<tr>
<th></th>
<th>Team IQ</th>
<th>Ave. Member IQ</th>
<th>Ave. Member IAOT</th>
<th>Team CRT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ave. Member IQ</td>
<td>0.39*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ave. Member IAOT</td>
<td>0.18</td>
<td>0.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Team CRT</td>
<td>0.24</td>
<td>0.41*</td>
<td>0.27</td>
<td></td>
</tr>
</tbody>
</table>
Group AOT and relative accuracy

```
## Call:
## lm(formula = bs_std ~ gaomt, data = ts)
##
## Residuals:
##    Min     1Q    Median     3Q    Max
## -0.9530 -0.2652 -0.0532  0.1963  2.1095
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.6939     0.0718   9.66  <2e-16 ***
## gaomt            -0.2022     0.0230  -8.80  <2e-16 ***
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 0.428 on 665 degrees of freedom
## (437 observations deleted due to missingness)
## Multiple R-squared:  0.104, Adjusted R-squared:  0.103
## F-statistic: 77.5 on 1 and 665 DF,  p-value: <2e-16
```

Each increment in GAOT related to ~0.2 SD increase in relative accuracy
Individual and group AOT

The mean of a team’s individual AOT scores explains no variance in relative performance when regressed with the group AOT score.

```
Call:
  lm(formula = bs_std ~ openminded + aomt.score, data = ts.ctt)

Residuals:
          Min          1Q       Median          3Q         Max
-0.5540 -0.1128  -0.0291  0.1595  0.7360

Coefficients:     Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.65595     1.15437    0.568    0.57258
openminded -0.20673     0.04887   -4.230    0.000107 ***
aomt.score  0.01022     0.19500    0.052    0.958432
---
Signif. codes:  < 0.001 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2475 on 47 degrees of freedom
Multiple R-squared:  0.2759,  Adjusted R-squared:  0.2451
F-statistic: 8.954 on 2 and 47 DF,  p-value: 0.0005073
```
Individual AOT is correlated with relative accuracy
Questions

- How to present discourse data? (Control, lasso regression, other?)
- Present turn-taking data? (Collinearity issues)
- Remove slopes photo?
- Integrative complexity?