The Long-Run Impacts of Psychotherapy on Depression, Beliefs, and Preferences

Bhargav Bhat    Jonathan de Quidt    Johannes Haushofer    Vikram Patel
Gautam Rao    Frank Schilbach    Pierre-Luc Vautrey

March 2022
Motivation

- Depression is common and costly
  - 20% life-time prevalence (Kessler & Wang 2009)
  - Often a recurrent condition

- Large treatment gaps, especially in developing countries (Chisholm et al. 2016)
  - Very limited supply of psychiatrists
  - 85% of Indians with major depressive disorders go untreated (Gururaj et al. 2016)

- Potential solution: simplified psychotherapy by non-specialist counselors
  - Effective in the short run (Singla et al. 2017; Barker et al. 2021)
This paper: follow-up study of two psychotherapy trials in Goa, India.

1. Long-run treatment effects on mental health
2. Do patients correctly predict the effects of therapy?
3. Impacts on self-confidence, belief formation
4. Impact on economic preferences and well-being (consumption and employment)
Follow-up of two RCTs in Goa, India

- Participants screened for depression using PHQ-9, a validated screening tool
Follow-up of two RCTs in Goa, India

- Participants screened for depression using PHQ-9, a validated screening tool

- **Treatment group**
  - Inexpensive therapies, delivered by non-specialists (cost ≈ $70 per patient)
  - Psychotherapy for depression, based mostly on behavioral activation
  - Relation between activity & mood; activity structuring & scheduling; problem solving
    - → Take actions to improve mood and reduce depression
    - → Taught tools and strategies to help deal with and prevent depression
Follow-up of two RCTs in Goa, India

• Participants screened for depression using PHQ-9, a validated screening tool

• **Treatment group**
  • Inexpensive therapies, delivered by non-specialists (cost ≈ $70 per patient)
  • Psychotherapy for depression, based mostly on behavioral activation
  • Relation between activity & mood; activity structuring & scheduling; problem solving
    → Take actions to improve mood and reduce depression
    → Taught tools and strategies to help deal with and prevent depression

• **Control group**
  • 'Treatment as usual': routine consultation with physician and/or gynaecologist
  • Providing PHQ-9 screening results to both the physician and the patient
Details of the two trials

- **Trial 1: “Healthy Activity Program (HAP)”** (2013-15, N=495)
  - Adults aged 18-65 years, excluding pregnant women
  - Screened on PHQ-9 score $\geq 15$ (moderately severe depression)
  - 6–8 sessions of therapy over 2-3 months
  - Delivered by individuals trained for 3 weeks
  - Approximately 70% completed treatment

- **Trial 2: “Thinking Healthy Program Peer-Delivered (THPP)”** (2015-17, N=280)
  - Pregnant women only, screened on PHQ-9 score $\geq 10$ (moderate depression)
  - 6–14 sessions of therapy over 7-12 months
  - Less involved version of intervention studied by Baranov et al. 2020 in Pakistan
Sample

- 75% follow-up rate across the two trials
  - Follow-up rates balanced across treatment/control
  - Covariates balanced across treatment/control
Sample

- 75% follow-up rate across the two trials
  - Follow-up rates balanced across treatment/control
  - Covariates balanced across treatment/control

- Pooled sample characteristics
  - 88% of total sample is female
  - 6 to 7 years of education on average
  - 1/3 employed at baseline
Sizable short-run effects of HAP → largely sustained for 5 years
Treatment effects on depression relative to existing studies

- Large average short-run treatment effects (0.46 SDs)
- Few long-run follow-up studies (except: Baranov et al. 2020)
- 91% of experts in economics and mental health who were asked to forecast these results underestimated the long-run effects.
Mechanism for persistent effect in case of HAP?

- Continued use of therapy in treatment group? Lack of availability
- Improvements in consumption/employment? No effects
Mechanism for persistent effect in case of HAP?

• Continued use of therapy in treatment group? Lack of availability

• Improvements in consumption/employment? No effects

• Possibility: Treated participants learn the principles or tools of behavioral activation + employ them to deal with future stresses

• Mediation analysis:
  • Short-run improvement in mental health is a strong mediator
  • Also role for extent of short-run behavioral activation
Do people understand treatment effectiveness?

- Seemingly lower-than-expected demand for therapy worldwide
  - Under-use in rich countries e.g. Cronin et al. (2021)
  - Surveys in 13 countries: lack of familiarity and confidence in therapy (Sapiens Lab 2021)

- Do people think therapy is effective? Does experiencing treatment change beliefs?

- We elicited people’s beliefs at endline about the treatment effects.
Experiencing treatment increases perceived effectiveness of HAP

- Control group underestimates the persistent effects of the Healthy Activity Program.
- Experiencing treatment corrects beliefs about long-run effects.
- No effects on short-run beliefs
How does therapy affect beliefs about oneself?

• We study the causal effect of psychotherapy on self-confidence

• And how self-confidence evolves in response to feedback
  • Optimistic belief-updating in response to feedback
    Eil & Rao 2011; Mobius et al. 2014; Zimmermann 2020

• Alternative hypotheses:
  (1) ‘Sadder but wiser’: Treating depression generates more overconfidence
    Korn et al. 2014; Alloy & Abramson 1979
  (2) ‘Protective optimism’: Therapy → more accurate view of self; less need for over-confidence
    Dunning and Story 1991; Blanton et al. 2001; Sherman and Cohen 2006
Paradigm, adapted from Möbius et al. (2021)

(1) Participants perform a “self-image relevant” task
   • Making bracelets – mimics realistic jobs

(2) Elicit prior on relative performance
   • Probability of above-median performance

(3) Provide noisy signal of truth

(4) Repeat ...

Benchmark: Bayes’ rule
Overconfidence in the Control Group

- Overconfidence: beliefs compared to full-information benchmarks
- Before feedback, control group is overconfident by 13 ppts.
- Feedback further increases initial overconfidence.
- Implies optimistic updating
Psychotherapy causes people to update less optimistically

- Treatment group updates their beliefs less optimistically
- Final beliefs are significantly less overconfident than control group’s.
- Suggest that therapy makes people “happier and wiser”
Belief updating relative to Bayesian benchmark

- Control group is close to Bayesian for positive signals; entirely ignores negative signals.
- Implies over-optimistic belief updating
Belief updating relative to Bayesian benchmark

- Control group is close to Bayesian for positive signals; entirely ignores positive signals.
- Implies over-optimistic belief updating
- Treatment group reacts less to positive signals; also ignores negative signals.
- Implies reduced asymmetry in updating in the treatment group
Impacts on self-confidence: discussion

- Therapy made people *less* overconfident – people seem “happier AND wiser”.
Impacts on self-confidence: discussion

- Therapy made people less overconfident – people seem “happier AND wiser”.

- Changes in depression or mood may not be underlying mechanisms
  - Similar finding in Trial 2, despite no long-run treatment effect on depression

- Suggests direct effects of therapy itself
  - May help people see themselves and the world more realistically
  - Makes some beliefs less ego-relevant
  - Help get better at reacting to feedback evenhandedly
Discussion

• Brief psychotherapy for depression can improve mental health for 5 years
  • Inexpensive (about $70 per person) and scalable intervention!
  • Averted 9.1 months of depression over 5 years at $7.25 per month averted
Discussion

• Brief psychotherapy for depression can improve mental health for 5 years
  • Inexpensive (about $70 per person) and scalable intervention!
  • Averted 9.1 months of depression over 5 years at $7.25 per month averted

• Therapy had lasting effects on people’s beliefs about themselves
  • Made people happier and more realistic about themselves
Discussion

• Brief psychotherapy for depression can improve mental health for 5 years
  • Inexpensive (about $70 per person) and scalable intervention!
  • Averted 9.1 months of depression over 5 years at $7.25 per month averted

• Therapy had lasting effects on people’s beliefs about themselves
  • Made people happier and more realistic about themselves

• Experiencing therapy makes people more positive about its efficacy.
  • May ex-ante underestimate effective treatments
  • Potential case for information interventions
Discussion

• Brief psychotherapy for depression can improve mental health for 5 years
  • Inexpensive (about $70 per person) and scalable intervention!
  • Averted 9.1 months of depression over 5 years at $7.25 per month averted

• Therapy had lasting effects on people’s beliefs about themselves
  • Made people happier and more realistic about themselves

• Experiencing therapy makes people more positive about its efficacy.
  • May ex-ante underestimate effective treatments
  • Potential case for information interventions

• Other results: increases in self-evaluated patience and altruism; no effects on employment or consumption
Thank you!

We’d love to hear your thoughts and feedback!

grao@fas.harvard.edu
Beliefs Updating Relative to Bayesian Benchmark

Round 1

Bayesian Benchmark

- Control
- Treatment

Coefficients

- Conservatism ($\beta$)
- Positive Signal ($\beta_H$)
- Negative Signal ($\beta_L$)
- Asymmetry ($\beta_H - \beta_L$)
Beliefs Updating Relative to Bayesian Benchmark

Round 2

Bayesian Benchmark

Conservatism ($\beta$)
Positive Signal ($\beta_H$)
Negative Signal ($\beta_L$)
Asymmetry ($\beta_H - \beta_L$)

Control
Treatment
Beliefs Updating Relative to Bayesian Benchmark

Appendix
Beliefs Updating Relative to Bayesian Benchmark

Round 4

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Control</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conservatism ($\beta$)</td>
<td>![Conservatism Bar Chart]</td>
<td>![Conservatism Bar Chart]</td>
</tr>
<tr>
<td>Positive Signal ($\beta_H$)</td>
<td>![Positive Signal Bar Chart]</td>
<td>![Positive Signal Bar Chart]</td>
</tr>
<tr>
<td>Negative Signal ($\beta_L$)</td>
<td>![Negative Signal Bar Chart]</td>
<td>![Negative Signal Bar Chart]</td>
</tr>
<tr>
<td>Asymmetry ($\beta_H - \beta_L$)</td>
<td>![Asymmetry Bar Chart]</td>
<td>![Asymmetry Bar Chart]</td>
</tr>
</tbody>
</table>

Bayesian Benchmark

-0.2 0 0.2 0.4 0.6 0.8 1
Appendix

Beliefs Updating Relative to Bayesian Benchmark

Round 5

Coeficients

- Conservatism ($\beta$)
- Positive Signal ($\beta_H$)
- Negative Signal ($\beta_L$)
- Asymmetry ($\beta_H - \beta_L$)

Bayesian Benchmark

Control

Treatment

Control

Treatment

Conservatism

Positive Signal

Negative Signal

Asymmetry

($\beta_H - \beta_L$)
Beliefs Updating Relative to Bayesian Benchmark

Pooled Rounds

Bayesian Benchmark
Control
Treatment

Conservatism ($\beta$)
Positive Signal ($\beta_H$)
Negative Signal ($\beta_L$)
Asymmetry ($\beta_H - \beta_L$)

Coefficients

Appendix
Defining Overconfidence

- Full-information benchmark
  - Suppose participants know their rank in the whole distribution of participants
  - Benchmark prior: probability of being in top half of group of 10 given rank
  - Benchmark posteriors: Bayesian updates from benchmark prior given signals

- Overconfidence: difference between elicited beliefs and benchmark
Beliefs Updating Relative to Bayesian Benchmark

Bayesian Benchmark

- Control
- Treatment

Coefficients
- Conservatism ($\beta$)
- Positive Signal ($\beta_H$)
- Negative Signal ($\beta_L$)
- Asymmetry ($\beta_H - \beta_L$)
Beliefs Updating Relative to Bayesian Benchmark

Appendix

Diagram showing coefficients for Conservatism, Positive Signal, Negative Signal, and Asymmetry, comparing Control and Treatment conditions to Bayesian Benchmark.
## Appendix

### Beliefs Updating Relative to Bayesian Benchmark

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Control</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Conservatism</strong> $(\beta)$</td>
<td><img src="#" alt="Control" /></td>
<td><img src="#" alt="Treatment" /></td>
</tr>
<tr>
<td><strong>Positive Signal</strong> $(\beta_H)$</td>
<td><img src="#" alt="Control" /></td>
<td><img src="#" alt="Treatment" /></td>
</tr>
<tr>
<td><strong>Negative Signal</strong> $(\beta_L)$</td>
<td><img src="#" alt="Control" /></td>
<td><img src="#" alt="Treatment" /></td>
</tr>
<tr>
<td><strong>Asymmetry</strong> $(\beta_H - \beta_L)$</td>
<td><img src="#" alt="Control" /></td>
<td><img src="#" alt="Treatment" /></td>
</tr>
</tbody>
</table>

### Diagram

The diagram above illustrates the changes in beliefs updating relative to the Bayesian benchmark. The bars represent the coefficients for different categories: Conservatism, Positive Signal, Negative Signal, and Asymmetry. The two bars for each category represent the Control and Treatment conditions, respectively.

The Bayesian Benchmark is shown as a dashed red line for reference, indicating the expected benchmark values for comparison.
# Impact on depression (PHQ9) and Mental Health with controls

<table>
<thead>
<tr>
<th></th>
<th>Full Sample</th>
<th>HAP</th>
<th>THPP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control mean (S.D.)</td>
<td>Treatment Effect (S.E.)</td>
<td>Control mean (S.D.)</td>
</tr>
<tr>
<td><strong>OLS without controls</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHQ-9 Score</td>
<td>7.97 (5.86)</td>
<td>-0.97** (0.48)</td>
<td>9.10 (5.97)</td>
</tr>
<tr>
<td>PHQ-9 ≥ 10</td>
<td>0.37 (0.48)</td>
<td>-0.08** (0.04)</td>
<td>0.45 (0.50)</td>
</tr>
<tr>
<td>Mood Score</td>
<td>6.49 (2.35)</td>
<td>0.38** (0.19)</td>
<td>6.19 (2.40)</td>
</tr>
<tr>
<td><strong>DML</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHQ-9 Score</td>
<td>7.97 (5.86)</td>
<td>-0.75 (0.48)</td>
<td>9.10 (5.97)</td>
</tr>
<tr>
<td>PHQ-9 ≥ 10</td>
<td>0.37 (0.48)</td>
<td>-0.07** (0.04)</td>
<td>0.45 (0.50)</td>
</tr>
<tr>
<td>Mood Score</td>
<td>6.49 (2.35)</td>
<td>0.38* (0.19)</td>
<td>6.19 (2.40)</td>
</tr>
</tbody>
</table>

| N                | 589                      | 395                      | 194                       |
Appendix

Healthy Activity Program: Effects on Economic Outcomes

Healthy Activity Program

<table>
<thead>
<tr>
<th>CONSUMPTION</th>
<th>WORK</th>
<th>CHILD EDUCATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.4</td>
<td>-0.3</td>
<td>-0.2</td>
</tr>
<tr>
<td>-0.1</td>
<td>0.0</td>
<td>0.1</td>
</tr>
<tr>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Pooled Treatment Effects (SDs)

Total

Consumption

Food

Durables

Other

Employed

Work Hours

Earnings

Reservation Wage

Years

Literacy

Work Hours

Earnings

Reservation Wage

Years

Literacy

Pooled Treatment Effects (SDs)
Appendix

Healthy Activity Program: Effects on Non-Economic Outcomes

Pooled Treatment Effects (SDs)

Healthy Activity Program

- **PREFERENCES**
- **IPV**
- **SLEEP**

<table>
<thead>
<tr>
<th>Time Index</th>
<th>Social Index</th>
<th>Risk Index</th>
<th>IPV Index</th>
<th>Beaten</th>
<th>Forced Sex</th>
<th>Sleep Index</th>
<th>Time Asleep</th>
<th>Sleep Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>−0.4</td>
<td>0.3</td>
<td>0.2</td>
<td>−0.1</td>
<td>0.0</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
</tr>
</tbody>
</table>

25 / 18
Appendix

Thinking Healthy Program: Effects on Economic Outcomes

Thinking Healthy Program

CONSUMPTION     WORK       CHILD EDUCATION

-0.4
-0.3
-0.2
-0.1
0.0
0.1
0.2
0.3
0.4
Total Total Total

Food Food Food

Durables Durables Durables

Other Other Other

Employed Work Hours Earnings Reservation Wage

Work Hours

Earnings Reservation Wage

Reservation Wage

Years

Years

Years

Literacy

Years

Literacy

Pooled Treatment Effects (SDs)

Thinking Healthy Program
Thinking Healthy Program: Effects on Economic Outcomes

Pooled Treatment Effects (SDs)

Thinking Healthy Program

- PREFERENCES
- IPV
- SLEEP
## Appendix

### Savings and Patience

<table>
<thead>
<tr>
<th></th>
<th>Saving Task (1)</th>
<th>Time Survey 1 (2)</th>
<th>Time Survey 2 (3)</th>
<th>Index (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment Dummy</td>
<td>0.076* (0.0383)</td>
<td>0.623** (0.23)</td>
<td>0.245 (0.21)</td>
<td>0.154** (0.0553)</td>
</tr>
<tr>
<td></td>
<td>[0.0425]</td>
<td>[0.253]</td>
<td>[0.234]</td>
<td>[0.0605]</td>
</tr>
<tr>
<td>Control Mean</td>
<td>0.723</td>
<td>7.33</td>
<td>8.26</td>
<td>0</td>
</tr>
<tr>
<td>Control SD</td>
<td>0.449</td>
<td>2.7</td>
<td>2.41</td>
<td>0.631</td>
</tr>
<tr>
<td>PIDs</td>
<td>487</td>
<td>494</td>
<td>494</td>
<td>486</td>
</tr>
</tbody>
</table>

**Figure 28/18**
### Social Preferences: Altruism

<table>
<thead>
<tr>
<th></th>
<th>Dictator Game (1)</th>
<th>Social Survey (2)</th>
<th>Index (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment Dummy</td>
<td>0.382 (0.959)</td>
<td>0.491** (0.214)</td>
<td>0.116* (0.06)</td>
</tr>
<tr>
<td>Control Mean</td>
<td>16.8</td>
<td>7.62</td>
<td>0</td>
</tr>
<tr>
<td>Control SD</td>
<td>11.2</td>
<td>2.49</td>
<td>0.691</td>
</tr>
<tr>
<td>PIDs</td>
<td>494</td>
<td>494</td>
<td>494</td>
</tr>
</tbody>
</table>
### Risk Aversion: Unwillingness to Accept Risky Bets/Investments

<table>
<thead>
<tr>
<th>Treatment Dummy</th>
<th>Risk Lottery (1)</th>
<th>Loss Lottery (2)</th>
<th>Risk Survey (3)</th>
<th>Index (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.425</td>
<td>4.02</td>
<td>-0.183</td>
<td>-0.00283</td>
</tr>
<tr>
<td></td>
<td>(2.82)</td>
<td>(5.76)</td>
<td>(0.243)</td>
<td>(0.0659)</td>
</tr>
<tr>
<td></td>
<td>[3.1]</td>
<td>[6.34]</td>
<td>[0.269]</td>
<td>[0.0735]</td>
</tr>
<tr>
<td>Control Mean</td>
<td>52.9</td>
<td>74.7</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Control SD</td>
<td>32.1</td>
<td>67</td>
<td>2.61</td>
<td>0.717</td>
</tr>
<tr>
<td>PIDs</td>
<td>494</td>
<td>494</td>
<td>494</td>
<td>494</td>
</tr>
</tbody>
</table>
## Death

<table>
<thead>
<tr>
<th></th>
<th>Whole (1)</th>
<th>With Suicide Thought (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment Dummy</td>
<td>-0.0252*</td>
<td>-0.0305</td>
</tr>
<tr>
<td></td>
<td>(0.0121)</td>
<td>(0.0179)</td>
</tr>
<tr>
<td></td>
<td>[0.0135]</td>
<td>[0.0196]</td>
</tr>
<tr>
<td>Control Mean</td>
<td>0.0393</td>
<td>0.0489</td>
</tr>
<tr>
<td>Control SD</td>
<td>0.195</td>
<td>0.216</td>
</tr>
<tr>
<td>PIDs</td>
<td>660</td>
<td>370</td>
</tr>
</tbody>
</table>
## Minimum Wage Required to Accept a Temporary Take-Home Job

**Table:** Reservation Wage to Accept a Flexible Take-Home Job

<table>
<thead>
<tr>
<th></th>
<th>Whole Sample</th>
<th>Premium Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Treatment dummy</td>
<td>71.9</td>
<td>47.4</td>
</tr>
<tr>
<td></td>
<td>(86.2)</td>
<td>(92.1)</td>
</tr>
<tr>
<td>Controls</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Mean ATE</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>Median ATE</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>Control Mean</td>
<td>1094</td>
<td>1094</td>
</tr>
<tr>
<td>Control SD</td>
<td>787</td>
<td>787</td>
</tr>
<tr>
<td>PIDs</td>
<td>340</td>
<td>340</td>
</tr>
</tbody>
</table>

- Job: making 1000 bracelets at home in 1 month
Appendix

Beliefs

- Task measures participants beliefs about their performance at a simple production task (bracelet-making using beads)
- First work on task for fixed time. Then we measure how good (productive) participants think they are relative to others.
  - Indicate beliefs about being in top versus bottom half out of 10 people by dividing water between two beakers
  - Get bonus for accurate guess
- Then, participants are given a “signal” about their performance relative to others and an opportunity to update their beliefs
- Participants’ beliefs are defined as below:

\[
\text{Belief}_{it} = \frac{\text{Water in upper beaker}_{it}}{\text{Water in upper beaker}_{it} + \text{Water in lower beaker}_{it}}
\]  

(1)
Beliefs

Belief updates regression equation is as below,

\[
\text{logit}(\text{Belief}_{it}) = \delta \text{logit}(\text{Belief}_{i,t-1}) + \beta_H I(s_{it} = H)\lambda_H + \beta_L I(s_{it} = L)\lambda_L + \\
\delta_d \text{logit}(\text{Belief}_{i,t-1}) \ast \text{PHQ9}_i + \beta_{H,d} I(s_{it} = H)\lambda_H \ast \text{PHQ9}_i + \\
\beta_{L,d} I(s_{it} = L)\lambda_L \ast \text{PHQ9}_i + \epsilon_{it}
\]  

(2)
Beliefs

- $Belief_{it}$ is posterior belief, while $Belief_{i,t-1}$ is prior belief, and $\text{logit}(p) = \log(p/1-p)$. Posterior belief in one round will become prior belief in the next round.
- $PHQ9_i$ is participants’ PHQ9 score in our data.
- $I(s_{it} = H)$ is an indicator for whether the signal $t$ is upper half, while $I(s_{it} = L)$ is an indicator for whether the signal $t$ is lower half.
- $\lambda_H$ is the log likelihood ratio of an upper signal if participants are in upper half, in our experiment, $\lambda_H$ is equal to $\ln 2$, on the contrary, $\lambda_L$ is the log likelihood ratio of a low signal if participants are in upper half, in our experiment, $\lambda_L$ is equal to $-\ln 2$.
- We don’t have to include a constant, because $I(s_{it} = H) + I(s_{it} = L) = 1$. 

### Table: Belief Updates (Interact with PHQ9 Score)

<table>
<thead>
<tr>
<th>Regressor</th>
<th>Round 1 (1)</th>
<th>Round 2 (2)</th>
<th>Round 3 (3)</th>
<th>Round 4 (4)</th>
<th>Round 5 (5)</th>
<th>All rounds (6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta$</td>
<td>0.80+++</td>
<td>0.48+++</td>
<td>0.53+++</td>
<td>0.75+++</td>
<td>0.64+++</td>
<td>0.62+++</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.09)</td>
<td>(0.09)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>$\beta_H$</td>
<td>0.58+++</td>
<td>0.67+++</td>
<td>0.67+</td>
<td>0.53+++</td>
<td>0.79</td>
<td>0.66+++</td>
</tr>
<tr>
<td></td>
<td>(0.18)</td>
<td>(0.17)</td>
<td>(0.18)</td>
<td>(0.16)</td>
<td>(0.19)</td>
<td>(0.10)</td>
</tr>
<tr>
<td>$\beta_L$</td>
<td>-0.04+++</td>
<td>0.00+++</td>
<td>-0.31+++</td>
<td>-0.14+++</td>
<td>-0.10+++</td>
<td>-0.13+++</td>
</tr>
<tr>
<td></td>
<td>(0.14)</td>
<td>(0.15)</td>
<td>(0.16)</td>
<td>(0.19)</td>
<td>(0.17)</td>
<td>(0.09)</td>
</tr>
<tr>
<td>$\delta_d$</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.01</td>
<td>-0.03**</td>
<td>-0.00</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>$\beta_{H,d}$</td>
<td>0.01</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>-0.02</td>
<td>-0.00</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>$\beta_{L,d}$</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.01</td>
<td>0.02</td>
<td>-0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
</tbody>
</table>

$P(\beta_H = \beta_L)$          | 0.01       | 0.01       | 0.00       | 0.02       | 0.00       | 0.00           |
$P(\beta_{H,d} = \beta_{L,d})$  | 0.69       | 0.71       | 0.62       | 0.58       | 0.44       | 0.62           |
$N$                                  | 359        | 355        | 354        | 352        | 350        | 1770           |
$R^2$                                | 0.54       | 0.46       | 0.52       | 0.45       | 0.54       | 0.49           |
Preliminary (non-causal) findings confirm depressive realism:

- Everyone is overconfident and uses feedback in an optimistic way (react more to good news than bad news)
- A higher PHQ9 is associated with a more moderate optimism in beliefs updating
## Belief updating results with different specifications

<table>
<thead>
<tr>
<th>Regressor</th>
<th>All rounds (1)</th>
<th>All rounds (2)</th>
<th>All rounds (3)</th>
<th>All rounds (4)</th>
<th>All rounds (5)</th>
<th>All rounds (6)</th>
<th>All rounds (7)</th>
<th>All rounds (8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{H,c}$</td>
<td>0.73 (0.08)</td>
<td>0.67 (0.08)</td>
<td>1.10 (0.11)</td>
<td>0.82 (0.09)</td>
<td>0.89 (0.08)</td>
<td>0.62 (0.08)</td>
<td>1.51 (0.09)</td>
<td>0.76 (0.08)</td>
</tr>
<tr>
<td>$\beta_{L,c}$</td>
<td>$-0.09$ (0.06)</td>
<td>$-0.08$ (0.06)</td>
<td>$-0.18$ (0.09)</td>
<td>$-0.12$ (0.07)</td>
<td>$0.35$ (0.04)</td>
<td>$0.27$ (0.04)</td>
<td>$0.78$ (0.07)</td>
<td>$0.33$ (0.05)</td>
</tr>
<tr>
<td>$\beta_{H,t}$</td>
<td>0.51 (0.06)</td>
<td>0.52 (0.06)</td>
<td>0.74 (0.09)</td>
<td>0.54 (0.07)</td>
<td>0.71 (0.05)</td>
<td>0.53 (0.04)</td>
<td>1.21 (0.07)</td>
<td>0.62 (0.07)</td>
</tr>
<tr>
<td>$\beta_{L,t}$</td>
<td>$-0.04$ (0.06)</td>
<td>$-0.06$ (0.06)</td>
<td>$-0.04$ (0.09)</td>
<td>$-0.03$ (0.07)</td>
<td>$0.47$ (0.05)</td>
<td>$0.30$ (0.06)</td>
<td>$0.90$ (0.08)</td>
<td>$0.38$ (0.07)</td>
</tr>
<tr>
<td>$P((\beta_{H,c} - \beta_{L,c})=(\beta_{H,t} - \beta_{L,t}))$</td>
<td>0.08</td>
<td>0.26</td>
<td>0.02</td>
<td>0.03</td>
<td>0.01</td>
<td>0.33</td>
<td>0.01</td>
<td>0.24</td>
</tr>
<tr>
<td>$P(\beta_{H,c} = \beta_{H,t})$</td>
<td>0.03</td>
<td>0.13</td>
<td>0.01</td>
<td>0.01</td>
<td>0.06</td>
<td>0.44</td>
<td>0.01</td>
<td>0.28</td>
</tr>
<tr>
<td>$P(\beta_{L,c} = \beta_{L,t})$</td>
<td>0.51</td>
<td>0.81</td>
<td>0.28</td>
<td>0.32</td>
<td>0.10</td>
<td>0.62</td>
<td>0.23</td>
<td>0.64</td>
</tr>
<tr>
<td>N</td>
<td>2620</td>
<td>2495</td>
<td>1563</td>
<td>2355</td>
<td>2123</td>
<td>1214</td>
<td>1066</td>
<td>915</td>
</tr>
</tbody>
</table>

- Drop updates with a degenerate prior or posterior: Yes
- Drop people with at least once degenerate belief: Yes
- Drop no updates: Yes
- Drop people never update their beliefs: Yes
- Drop wrong updates: Yes
- Drop people with at least one wrong updates: Yes
Screening and enrollment statistics for Trial 1

- Total screened 31,888 PHC attendees
- 785 patients (2.5%) scored 15 or above on PHQ-9 and eligible.
- From these 495 patients (63.1%) consented for participating in the HAP Trial and 290 patients refused consent.
- The recruitment period was 18 months (extended from 1 year planned initially to get the sample size).
Appendix

Screening and enrollment statistics for Trial 2

- Total screened 6369 pregnant women
- 335 women (5%) scored 10 or above on PHQ-9 and eligible.
- From these 280 (84%) consented for participating in the HAP Trial and 55 women refused consent.
- The recruitment period was 20 months (extended from 18 months planned initially to get the sample size).
Healthy Activity Program: Details

- Light psychotherapy based on behavioral activation, a subset of Cognitive Behavioral Therapy (CBT)
- Delivered by lay counselors trained during 3 weeks of workshops and 6 months of a practice internship
- 6 to 8 sessions of 40 minutes
- Key Components
  - Behavioral assessment, self-monitoring
  - Psycho-education on relationship between activity and mood
  - Activity structuring and scheduling
  - Problem solving
Thinking Healthy Program: Details

- Light psychotherapy based mostly on behavioral activation.
  - Identification and planning of pleasurable activities and problem solving
  - Education about healthy eating and sleeping
- Delivered by a “peer”: typically slightly older woman who has had at least one child, living in the same community
- Before childbirth, around once every 2 weeks. After childbirth, around once every 2-3 months. Between 6 and 14 sessions in total.
- Stronger concerns about exclusion restriction (more below).
Appendix

Survey of Experts

<table>
<thead>
<tr>
<th>Variable</th>
<th>Prediction IQR</th>
<th>Actual pooled estimate</th>
<th>HAP estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHQ9 Score</td>
<td>-0.2</td>
<td>0</td>
<td>-0.2</td>
</tr>
<tr>
<td>Initial Overconfidence</td>
<td>-0.2</td>
<td>0</td>
<td>-0.2</td>
</tr>
<tr>
<td>Patience index</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Risk tolerance index</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Altruism index</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Consumption</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Employment</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Pooled Treatment Effect (SDs)
### Survey of Experts: Belief Updating

<table>
<thead>
<tr>
<th>Belief updating ratio</th>
<th>Prediction IQR</th>
<th>Actual pooled estimate</th>
<th>HAP estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Belief updating ratio values: -0.5, 0, 0.5, 1, 1.5, 2.
### Treatment Effect Heterogeneity on PHQ-9 Score

<table>
<thead>
<tr>
<th></th>
<th>Base Model</th>
<th>X=Above Median Base PHQ-9</th>
<th>X=Above Median Age</th>
<th>X= Above Median Male</th>
<th>X= Above Median Education</th>
<th>X= Above Median Predicted Drop Expectations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Full Sample - Including Trial Indicators</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment Effect</td>
<td>-0.96</td>
<td>-1.83</td>
<td>0.43</td>
<td>-1.39</td>
<td>-1.13</td>
<td>-1.82 (0.81)</td>
</tr>
<tr>
<td></td>
<td>0.49</td>
<td>(0.86)</td>
<td>(1.42)</td>
<td>(0.70)</td>
<td>(0.64)</td>
<td>(0.81)</td>
</tr>
<tr>
<td>X</td>
<td>0.82</td>
<td>0.95</td>
<td>-3.23</td>
<td>-0.18</td>
<td>-1.57</td>
<td>-1.6 (0.65)</td>
</tr>
<tr>
<td></td>
<td>(0.73)</td>
<td>(1.04)</td>
<td>(0.96)</td>
<td>(0.80)</td>
<td>(0.64)</td>
<td>(0.65)</td>
</tr>
<tr>
<td>Treatment * X</td>
<td>0.48</td>
<td>-2.36</td>
<td>0.17</td>
<td>-1.52</td>
<td>0.84</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>(1.09)</td>
<td>(1.52)</td>
<td>(1.30)</td>
<td>(1.11)</td>
<td>(0.97)</td>
<td>(0.98)</td>
</tr>
<tr>
<td>THPP</td>
<td>-3.04</td>
<td>-2.70</td>
<td>-4.02</td>
<td>-3.37</td>
<td>-3.51</td>
<td>-3.42</td>
</tr>
<tr>
<td></td>
<td>(0.68)</td>
<td>(1.05)</td>
<td>(0.68)</td>
<td>(0.69)</td>
<td>(0.65)</td>
<td>(0.64)</td>
</tr>
<tr>
<td>Treat * THPP</td>
<td>1.65</td>
<td>-0.35</td>
<td>1.36</td>
<td>1.80</td>
<td>1.44</td>
<td>1.39</td>
</tr>
<tr>
<td></td>
<td>(1.08)</td>
<td>(1.55)</td>
<td>(1.05)</td>
<td>(1.04)</td>
<td>(1.00)</td>
<td>(0.99)</td>
</tr>
</tbody>
</table>
Appendix

Treatment Effect Heterogeneity on PHQ-9 Score, controlling trial arm

![Graph showing treatment effects on PHQ-9 score for different factors such as age, education, gender, baseline, and PHQ-9 remission and efficacy, with below and above median scores compared.](image-url)
## Impacts on PHQ-9 components

<table>
<thead>
<tr>
<th>PHQ Questions</th>
<th>Full Sample</th>
<th>HAP</th>
<th>THPP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control mean (S.D.)</td>
<td>Treatment Effect (S.E.)</td>
<td>Control mean (S.D.)</td>
</tr>
<tr>
<td>Sleeping Difficulty</td>
<td>1.22 (1.15)</td>
<td>-0.14 (0.09)</td>
<td>1.40 (1.15)</td>
</tr>
<tr>
<td>Tiredness</td>
<td>1.39 (1.08)</td>
<td>-0.13 (0.09)</td>
<td>1.52 (1.06)</td>
</tr>
<tr>
<td>Poor Appetite</td>
<td>0.89 (1.15)</td>
<td>-0.23** (0.09)</td>
<td>1.00 (1.19)</td>
</tr>
<tr>
<td>Trouble Concentrating</td>
<td>0.84 (1.12)</td>
<td>-0.04 (0.09)</td>
<td>0.96 (1.17)</td>
</tr>
<tr>
<td>Little Interest/Pleasure</td>
<td>0.77 (1.08)</td>
<td>0.06 (0.09)</td>
<td>0.90 (1.14)</td>
</tr>
<tr>
<td>Feeling Depressed</td>
<td>1.12 (1.15)</td>
<td>-0.14 (0.09)</td>
<td>1.28 (1.21)</td>
</tr>
<tr>
<td>Feeling Bad About Oneself</td>
<td>1.05 (1.19)</td>
<td>-0.24*** (0.09)</td>
<td>1.25 (1.22)</td>
</tr>
<tr>
<td>Abnormal Speech or Movement</td>
<td>0.35 (0.75)</td>
<td>-0.04 (0.06)</td>
<td>0.37 (0.74)</td>
</tr>
<tr>
<td>Better off Dead/Self Harm</td>
<td>0.33 (0.76)</td>
<td>0.00 (0.06)</td>
<td>0.41 (0.82)</td>
</tr>
</tbody>
</table>

N  
589 395 194
Very similar picture for average PHQ-9 scores

Healthy Activity Program (N = 493)

Thinking Healthy Program – Peer (N = 280)