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 \approx \$70 per patient)
- Psychotherapy for depression, based mostly on behavioral activation
- Relation between activity & mood; activity structuring & scheduling; problem solving
 - \rightarrow Take actions to improve mood and reduce depression
 - \rightarrow 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 pprox \$70 per patient)
- Psychotherapy for depression, based mostly on behavioral activation
- Relation between activity & mood; activity structuring & scheduling; problem solving
 - \rightarrow Take actions to improve mood and reduce depression
 - \rightarrow 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 \rightarrow largely sustained for 5 years



Panel A: Fraction depressed over time in HAP

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.

Depression

Experiencing treatment increases perceived effectiveness of HAP



X Treatment group's belief (90% CI)

- 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
 - Optimisic 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 (Details)
- 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"



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



- 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

- 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

- 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

- 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

- 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

Introduction

Thank you!

We'd love to hear your thoughts and feedback!

grao@fas.harvard.edu



Thinking Healthy Program (Control Group)







Round 2



Round 3



Beliefs Updating Relative to Bayesian Benchmark



Round 5



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







$\operatorname{Appendix}$

Impact on depression (PHQ9) and Mental Health with controls

	Full Sample			НАР	THPP	
	Control mean	Treatment Effect	Control mean	Treatment Effect	Control mean	Treatment Effect
	(S.D.)	(S.E.)	(S.D.)	(S.E.)	(S.D.)	(S.E.)
OLS without controls						
PHQ-9 Score	7.97	-0.97**	9.10	-1.43**	5.68	-0.04
	(5.86)	(0.48)	(5.97)	(0.61)	(4.92)	(0.78)
$PHQ-9 \ge 10$	0.37 (0.48)	-0.08** (0.04)	0.45 (0.50)	-0.12** (0.05)	0.22 (0.41)	-0.00 (0.06)
Mood Score	6.49	0.38**	6.19	0.38	7.10	0.40
	(2.35)	(0.19)	(2.40)	(0.24)	(2.14)	(0.32)
DML						
PHQ-9 Score	7.97	-0.75	9.10	-1.26**	5.68	0.19
	(5.86)	(0.48)	(5.97)	(0.60)	(4.92)	(0.77)
PHQ-9≥ 10	0.37	-0.07**	0.45	-0.12**	0.22	0.01
	(0.48)	(0.04)	(0.50)	(0.05)	(0.41)	(0.06)
Mood Score	6.49	0.38*	6.19	0.41*	7.10	0.34
	(2.35)	(0.19)	(2.40)	(0.24)	(2.14)	(0.31)
Ν	589			395	194	

Healthy Activity Program: Effects on Economic Outcomes

Healthy Activity Program



Healthy Activity Program: Effects on Non-Economic Outcomes

Healthy Activity Program



Thinking Healthy Program: Effects on Economic Outcomes

Thinking Healthy Program



Thinking Healthy Program: Effects on Economic Outcomes

Thinking Healthy Program



Savings and Patience

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

Figure

Social Preferences: Altruism

Dictator Game	Social Survey	Index
(1)	(2)	(3)
0.382	0.491**	0.116*
(0.959)	(0.214)	(0.06)
[1.09]	[0.237]	[0.0678]
16.8	7.62	0
11.2	2.49	0.691
494	494	494
	Dictator Game (1) 0.382 (0.959) [1.09] 16.8 11.2 494	Dictator GameSocial Survey(1)(2)0.3820.491**(0.959)(0.214)[1.09][0.237]16.87.6211.22.49494494

Figure

Risk Aversion: Unwillingness to Accept Risky Bets/Investments

	Risk Lottery	Loss Lottery	Risk Survey	Index
	(1)	(2)	(3)	(4)
Treatment Dummy	0.425	4.02	-0.183	-0.00283
	(2.82)	(5.76)	(0.243)	(0.0659)
	[3.1]	[6.34]	[0.269]	[0.0735]
Control Mean	52.9	74.7	7	0
Control SD	32.1	67	2.61	0.717
PIDs	494	494	494	494



Death

	Whole (1)	With Suicide Thought (2)
Treatment Dummy	-0.0252*	-0.0305
	(0.0121)	(0.0179)
	[0.0135]	[0.0196]
Control Mean	0.0393	0.0489
Control SD	0.195	0.216
PIDs	660	370



Minimum Wage Required to Accept a Temporary Take-Home Job

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

	Whole Sample			Premium Sample				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment dummy	71.9	47.4	115	116	98.4	71.8	144	145
	(86.2)	(92.1)	(87.4)	(87.7)	(105)	(118)	(105)	(105)
	. ,	. ,	[97]	[93.6]	. ,	. ,	[113]	[110]
Controls	NO	YES	DML	DML	NO	YES	DML	DML
Mean ATE	NO	NO	YES	NO	NO	NO	YES	NO
Median ATE	NO	NO	NO	YES	NO	NO	NO	YES
Control Mean	1094	1094	1094	1094	1080	1080	1080	1080
Control SD	787	787	787	787	812	812	812	812
PIDs	340	340	340	340	236	236	236	236

• Job: making 1000 bracelets at home in 1 month

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:

$$\mathsf{Belief}_{it} = \frac{\mathsf{Water in upper beaker}_{it}}{\mathsf{Water in upper beaker}_{it} + \mathsf{Water in lower beaker}_{it}} \tag{1}$$

Beliefs

Belief updates regression equation is as below,

$$logit(Belief_{it}) = \delta logit(Belief_{i,t-1}) + \beta_H I(s_{it} = H)\lambda_H + \beta_L I(s_{it} = L)\lambda_L + \\ \delta_d logit(Belief_{i,t-1}) * PHQ9_i + \beta_{H,d} I(s_{it} = H)\lambda_H * PHQ9_i + \\ \beta_{L,d} I(s_{it} = L)\lambda_L * PHQ9_i + \epsilon_{it}$$
(2)

Beliefs

- Belief_{it} is posterior belief, while Belief_{i,t-1} is prior belief, and logit(p) = log(p/1-p). Posterior belief in one round will become prior belief in the next round.
- *PHQ*9_{*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.
- λ_H is the log likelihood ratio of an upper signal if participants are in upper half, in our experiment, λ_H is equal to ln2, on the contrary, λ_L is the log likelihood ratio of a low signal if participants are in upper half, in our experiment, λ_L is equal to -ln2.
- We don't have to include a constant, because $I(s_{it} = H) + I(s_{it} = L) = 1$.

$\operatorname{Appendix}$

Beliefs

Regressor	Round 1 (1)	Round 2 (2)	Round 3 (3)	Round 4 (4)	Round 5 (5)	All rounds (6)
δ	0.80++	0.48+++	0.53+++	0.75++	0.64+++	0.62+++
	(0.09)	(0.09)	(0.09)	(0.11)	(0.11)	(0.06)
β_H	0.58^{++}	0.67++	0.67^{+}	0.53+++	0.79	0.66+++
	(0.18)	(0.17)	(0.18)	(0.16)	(0.19)	(0.10)
β_L	-0.04^{+++}	0.00+++	-0.31^{+++}	-0.14^{+++}	-0.10^{+++}	-0.13^{+++}
	(0.14)	(0.15)	(0.16)	(0.19)	(0.17)	(0.09)
δ_d	-0.01	-0.00	0.01	-0.03^{**}	-0.00	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$\beta_{H,d}$	0.01	0.00	-0.00	0.00	-0.02	-0.00
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)
$\beta_{L,d}$	0.00	-0.00	0.01	0.02	-0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
$P(\beta_H = \beta_L)$	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

Table: Belief Updates (Interact with PHQ9 Score)

Beliefs

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

Regressor	All rounds							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\beta_{H,c}$	0.73	0.67	1.10	0.82	0.89	0.62	1.51	0.76
	(0.08)	(0.08)	(0.11)	(0.09)	(0.08)	(0.08)	(0.08)	(0.09)
$\beta_{L,c}$	-0.09	-0.08	-0.18	-0.12	0.35	0.27	0.78	0.33
	(0.06)	(0.06)	(0.09)	(0.07)	(0.04)	(0.04)	(0.07)	(0.05)
$\beta_{H,t}$	0.51	0.52	0.74	0.54	0.71	0.53	1.21	0.62
	(0.06)	(0.06)	(0.08)	(0.06)	(0.06)	(0.08)	(0.07)	(0.09)
$\beta_{L,t}$	-0.04	-0.06	-0.04	-0.03	0.47	0.30	0.90	0.38
	(0.06)	(0.06)	(0.09)	(0.07)	(0.05)	(0.06)	(0.08)	(0.07)
$P((\beta_{H,c} - \beta_{L,c}) = (\beta_{H,t} - \beta_{L,t}))$	0.08	0.26	0.02	0.03	0.01	0.33	0.01	0.24
$P(\beta_{H,c} = \beta_{H,t})$	0.03	0.13	0.01	0.01	0.06	0.44	0.01	0.28
$P(\beta_{L,c} = \beta_{L,t})$	0.51	0.81	0.28	0.32	0.10	0.62	0.23	0.64
N	2620	2495	1563	2355	2123	1214	1066	915
Drop updates with a degenerate prior or posterior	Yes							
Drop people with at least once degenerate belief		Yes						Yes
Drop no updates			Yes				Yes	
Drop people never update their beliefs				Yes				Yes
Drop wrong updates					Yes	Yes	Yes	Yes
Drop people with at least one wrong updates						Yes		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).

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).





Prediction IQR^o Actual pooled estimate • HAP estimate

 $\operatorname{Appendix}$

Survey of Experts: Belief Updating



$\operatorname{Appendix}$

Treatment Effect Heterogeneity on PHQ-9 Score

	Base Model (S.E.)	X=Above Median Base PHQ-9 (S.E.)	X=Above Median Age (S.E.)	X= Male (S.E.)	X=Above Median Education (S.E.)	X=Above Median Predicted PHQ-9 Drop (S.E.)	X=Above Median Expectations (S.E.)
Full Sample - Including Trial Indicators							
Treatment Effect	-0.96	-1.83	0.43	-1.39	-1.13	-1.82	-1.54
	0.49	(0.86)	(1.42)	(0.70)	(0.64)	(0.81)	(0.81)
х		0.82	0.95	-3.23	-0.18	-1.57	-0.16
		(0.73)	(1.04)	(0.96)	(0.80)	(0.64)	(0.65)
Treatment * X		0.48	-2.36	0.17	-1.52	0.84	0.20
		(1.09)	(1.52)	(1.30)	(1.11)	(0.97)	(0.98)
THPP		-3.04	-2.70	-4.02	-3.37	-3.51	-3.42
		(0.68)	(1.05)	(0.68)	(0.69)	(0.65)	(0.64)
Treat * THPP		1.65	-0.35	1.36	1.80	1.44	1.39
		(1.08)	(1.55)	(1.05)	(1.04)	(1.00)	(0.99)
THPP Treat * THPP		-3.04 (0.68) 1.65 (1.08)	-2.70 (1.05) -0.35 (1.55)	-4.02 (0.68) 1.36 (1.05)	-3.37 (0.69) 1.80 (1.04)	-3.51 (0.65) 1.44 (1.00)	-3.42 (0.64) 1.39 (0.99)

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



Impacts on PHQ-9 components

Con	ontrol mean .D.)	Treatment Effect	Control mean	Treatment Effect	6	
(3.1		(3.L.)	(S.D.)	(S.E.)	Control mean (S.D.)	Treatment Effect (S.E.)
PHQ Questions						
Sleeping Difficulty 1.22 (1.1	.15)	-0.14 (0.09)	1.40 (1.15)	-0.16 (0.11)	0.86	-0.08 (0.15)
Tiredness 1.39 (1.0	39 .08)	-0.13	1.52	-0.17	1.12	-0.04
Poor Appetite 0.89 (1.1	39 .15)	-0.23** (0.09)	1.00 (1.19)	-0.35*** (0.11)	0.66 (1.02)	0.03 (0.15)
Trouble Concentrating 0.84 (1.1	34 · · · · · · · · · · · · · · · · · · ·	-0.04 (0.09)	0.96 (1.17)	-0.08 (0.11)	0.61 (1.00)	0.02 (0.14)
Little Interest/Pleasure 0.77 (1.0	77 .08)	0.06 (0.09)	0.90 (1.14)	0.01 (0.11)	0.50 (0.88)	0.12 (0.13)
Feeling Depressed 1.12 (1.1	12 .15)	-0.14 (0.09)	1.28 (1.21)	-0.21* (0.12)	0.79 (0.94)	-0.00 (0.13)
Feeling Bad About Onesself 1.05 (1.1	05 .19)	-0.24*** (0.09)	1.25 (1.22)	-0.38*** (0.11)	0.64 (1.01)	0.06 (0.15)
Abnormal Speech or Movement 0.35 (0.7	35 .75)	-0.04 (0.06)	0.37 (0.74)	-0.02 (0.08)	0.32 (0.79)	-0.08 (0.11)
Better off Dead/Self Harm 0.33 (0.7	33 .76)	0.00 (0.06)	0.41 (0.82)	-0.03 (0.08)	0.18 (0.59)	0.06 (0.09)
Ν	5	89	3	395		194

Very similar picture for average PHQ-9 scores

Healthy Activity Program (N = 493)



Thinking Healthy Program – Peer (N = 280)

