Discrimination and Depression: The role of implicit Theory of the Mind and Intergroup bias

By Eoin Williams

Supervisor: Dr. Sarah White (assisted by Yongxi Xiao MD)

Introduction to Research Problem

People with mental illness experience disproportionate stigmatization and discrimination (Thornicroft et al., 2007). This is a woefully common issue, with over 48% of mentally ill persons reporting being shunned or discriminated against at work (Brohan et al, 2023). A variety of factors are implicated in contributing to discrimination against the mentally ill, from subtle word use to the reactions of an entire society (Gronholm et al, 2024). Understanding and narrowing down what causes this may facilitate changes to halt such discrimination.

Intergroup bias is the tendency to favour ingroup members_– those considered to be similar to ourselves- adults are less likely to feel empathy towards outgroup members (Avenanti et al., 2010). This type of prejudice has long been linked to intergroup conflict and outgroup discrimination (Wilder, 1978). Implicit Theory of the Mind (iToM) refers to subconscious use of folk psychology and the knowledge that mental states differ between different people to consider the mental state of oneself and others (Quesque et al, 2024). IToM has been linked to intergroup bias (Molenberghs et al, 2018). Studies have shown that the desire to distance mentally ill persons from society has not changed in recent years (Henderson et al, 2019). It seems that society would be more likely to subconsciously discriminate against the mentally ill rather than consciously do so. Due to this, this study will examine each participant's implicit ToM (iToM) to study their subconscious.

IToM and, by extent, intergroup bias will be assessed via a true belief/ false belief (TB/ FB) eye-tracking paradigm. Recruitment and data analysis focused on individuals with depression as depressive traits can easily be assessed with questionnaires.

Ethical approval for this project was provided by the UCL Research Ethics Committe.

Methods

Demographics

Total participants:92

Sex:
Male- 22
Female- 70

Ethnicity:
Asian- 73
White- 13
Black- 5
Mixed- 1

Education:
PhD or Masters- 29
Bachalor's or Polytechnic- 37
Undergraduate- 17
Three or more A-levels- 8
Five or more GCSEs- 1

English familiararity:
First Language- 26
Second or third language- 66

Presence of mental illness:
Current depression diagnosis- 9
Previous depression diagnosis- 10
Anxiety or panic disorders- 4
Anorexia- 1
Comorbid depression- 3

Recruitment

The UCL Psychology Subject Pool (Sona Systems, n.d.) was used for recruitment with an advertisement being placed on this software for a "FUN VIDEO WATCHING EXPERIMENT (in person) !!". Participants, over 35 and under 18, and those with hearing or vision impairments were excluded. This was because visual or auditory impairments may reduce the reliability of the eye-tracking data, and to focus on young people with depression. Each participant was compensated £9 for their time and/ or travel expenses to arrive at the lab.

Materials

Tobii Fusion Pro eye-trackers were connected to Dell Latitude 5540 laptops which used windows 11 to run Tobii Pro Lab eye-tracking software version 1.171 (Tobii AB 2021).

Procedure and Data Collection

Eye tracking data was obtained through a TB/ FB paradigm created by Yongxi Xiao and based on a study by Southgate et al in 2007.

First, a vignette of Hazel or Sophie would be shown to the participant. If Hazel was shown, she would describe her life as happy and fulfilling. If Sophie was shown, she would describe her life as empty and overwhelming. The participant would then observe a familiarization block featuring the actor seen in the vignette and two experimental blocks. The familiarization block starts with one short video showing the actor moving her right hand through a flap to grab the ball on top of the right box after a chime is heard. The other short video is much the same but with her left hand reaching to the ball on top of the left box. This counterbalances the ball's location. Two long videos are then shown. These involve the puppet taking the ball from the center of the screen and placing it into either the left or the right box before the chime sounds, signaling the actor to reach through into the box and retrieve the ball. While only two long videos are shown, there are 4 possible videos the participant could see. One where the puppet places the ball into the left box and one into the right box. The other two possible videos are versions of the previously described videos but with the screen flipped. So, the left box appears on the right and the right box appears on the left. These trials are designed to train the participant to look at the flap they think the actor will reach into when the chime is heard.

Figure 1. *Sophie picking up the ball after the chime sounds during a long familiarization task.*

Each experimental block started with one short familiarization and one long familiarization before 8 TB tasks and 8 FB tasks are shown in a random order. All TB tasks involved the puppet moving the ball from the center of the screen into one of the boxes, then switching the ball's location into the other box and disappearing off-screen. A doorbell then rings, but the actor refuses to answer, and the puppet reappears to switch the ball's location a final time before disappearing for the chime to be heard. FB tasks are like the TB tasks, but this time, the actor turns to answer the doorbell. She, thus, does not see the puppet repaper to take the ball off-screen. Therefore, the actor has a false belief regarding the ball's location. The actor then turns back around for the chime to play. What sets different TB tasks apart and different FB tasks apart is variation in the handedness of the puppeteer, the box the ball is initially placed in and whether the video is flipped or not. Some videos had a left-handed puppeteer place the ball in the left box, some had a right-handed puppeteer place the ball in the right box, some had a left-handed puppeteer place the ball in the right box and vice versa. Each of these four possibilities had a version of itself with the video flipped over. This makes a total of 8 TB videos and 8 FB videos.

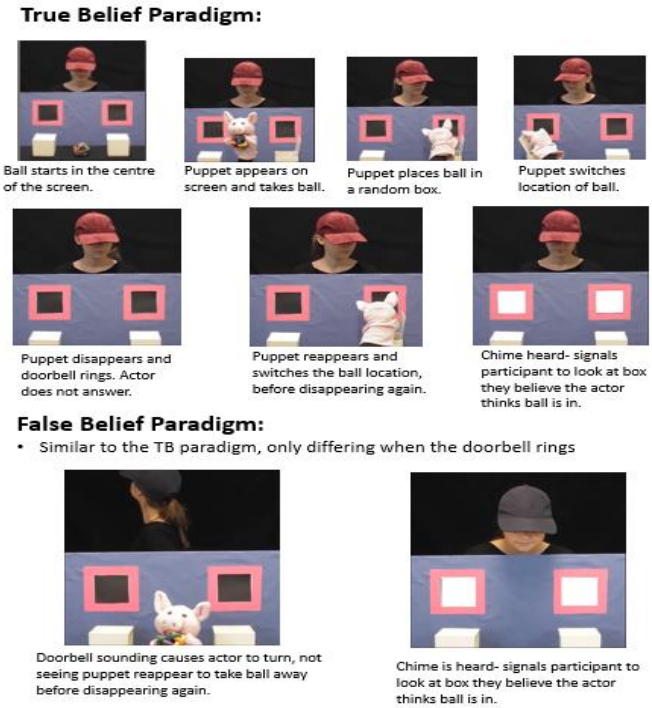

Figure 2. *Screenshots of the true belief and false belief paradigms. In this case, Hazel is in red and Sophie in grey.*

After the first experimental block is shown, the participant will be given a manipulation check questionnaire related to the actor they just viewed. The participant will then view the vignette and second experimental block containing the actor they have yet to be introduced to. The only difference between the two experimental blocks is that one contains Hazel,

and one contains Sophie. After the second experimental block is shown, the participant will be given another manipulation check questionnaire regarding the actor they just viewed. A questionnaire was then given to the participant to determine whether the participant identified with depressed or non- depressed persons. Once this was complete, the Beck Depression Inventory II (Beck, et al 1996) or BDI-II was used to assess the depressive symptoms of each participant. To further assess the prevalence of depression in the participants, the Patient Health Questionnaire-9 (Kroenke, et al 2001) or PHQ-9 was then completed by each subject. Following these assessments, each participant completed a demographics sheet and a feedback form.

At random, either Hazel would be wearing a red hat and Sophie a grey hat or Hazel would be wearing a red hat and Sophie a grey hat. This manipulation check assessed if hat colour impacted the participants iToM. The hat simultaneously increased the validity of the eye-tracking data by preventing the participant from following the actor's gaze.

Data Processing and Analysis

Of the 92 participants, 38 were excluded, leaving 54 participants for analysis. These people were excluded due to having incomplete eye-tracking data, being over 35 and/ or self-reporting that they assumed this was an experiment regarding iToM. These participants therefore used explicit ToM rather than iToM, compromising their results. The eye-tracking data was processed by calculating a Differential Looking Score (DLS) where the time spent looking at the flap the actor would not believe the ball to be behind was subtracted from the time looking at the flap the actor believed the ball to be behind. Therefore, the DLS score of a participant who looked at the correct flap for longer would be positive and the DLS score of a participant who looked at the incorrect flap for longer would be negative. SPSS (IBM, 2022) was then used to perform a 2 by 2 ANVOA to search for significant difference between mean DLS scores of either actor (Hazel and Sophie) and either belief (TB or FB). 2 by 2 ANCOVAs were then run looking for the same difference but with either BDI scores, PHQ-9 scores, the presence of depression or the relatability of each actor as covariates. A paired sample t-test was performed on actor relatability. Individual correlation analyses were also performed between DLS scores and each previously mentioned factor. All figures in results and discussion section were produced in SPSS (IBM, 2022).

Results and Discussion

Most of the ANOVAs and ANCOVAs run did not reach statistical significance ($p<0.05$). No significant difference was found between the DLS scores of Hazel or Sophie regardless of if they held a TB or FB, even when BDI-II and PHQ-9 scores were used as covariates. This provides no evidence for the possibility of intergroup bias resulting in discrimination due to inaccurate iToM against depressed persons.

One ANCOVA which did reach significance ($p=0.005$) was when the relatability of Hazel and Sophie was used as a covariate, uncovering a significant difference in DLS scores of true and false belief tasks in both Hazel's and Sophie's performances.

4. Actor * Belief

Measure: MEASURE_1

Actor	Belief	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	1	.106[a]	.044	.017	.195
	2	.154[a]	.036	.081	.227
2	1	.092[a]	.039	.014	.171
	2	.122[a]	.039	.043	.201

a. Covariates appearing in the model are evaluated at the following values: Hazel Relatability = 7.83, Sophie Relatability = 9.24.

Figure 3. *Mean TB and FB DLS scores for Hazel and Sophie. Actor 1=Sophie. Actor 2= Hazel. Belief 1=FB. Belief 2=TB.*
FB tasks produced significantly lower DLS scores than TB tasks. This may be because FB tasks require greater out of reality processing than TB tasks. In a FB task the actor's belief and the ball's location do not match and certain brain areas are activated to a greater degree during FB tasks (Döhnel et al, 2012). Using actor relatability as a covariate may have corrected for differences in DLS scores due to intergroup bias against unrelatable actors, rather than against actors with depression.

A paired sample t-test of the relatability of Hazel compared to Sophie revealed that participants found Sophie 1.407 points more relatable than Hazel on a scale of 1-15 ($p<0.001$). This suggests that most participants are more likely to consider Sophie part of an ingroup and Hazel part of an outgroup. None of the following correlation analysis reached $p<0.05$, but results are still of note.

Correlation analysis of the DLS scores from Hazel and Sophie's FB and TB tasks with Hazel and Sophie's relatability scores found a negative correlation between Hazel's relatability and DLS scores on Sophie's FB tasks as well as a negative correlation between Sophie's relatability and Hazel's TB tasks. Another negative correlation was found between DLS scores for Sophie's FB tasks and Hazel's FB tasks. This further implies that actor relatability rather than the presence of depression will produce an intergroup bias which worsens iToM ability.

Correlations

		DLS Score Sophie FB (seconds)	DLS Score Sophie TB (seconds)	DLS Score Hazel FB (seconds)	DLS Score Hazel TB (seconds)	Hazel Relatability	Sophie Relatability
DLS Score Sophie FB (seconds)	Pearson Correlation	1	.190	-.086	.076	-.248	.025
	Sig. (2-tailed)		.168	.536	.583	.071	.856
	N	54	54	54	54	54	54
DLS Score Sophie TB (seconds)	Pearson Correlation	.190	1	.043	.128	.065	.182
	Sig. (2-tailed)	.168		.755	.358	.642	.188
	N	54	54	54	54	54	54
DLS Score Hazel FB (seconds)	Pearson Correlation	-.086	.043	1	.046	.189	.238
	Sig. (2-tailed)	.536	.755		.741	.171	.083
	N	54	54	54	54	54	54
DLS Score Hazel TB (seconds)	Pearson Correlation	.076	.128	.046	1	.030	-.267
	Sig. (2-tailed)	.583	.358	.741		.830	.051
	N	54	54	54	54	54	54
Hazel Relatability	Pearson Correlation	-.248	.065	.189	.030	1	.122
	Sig. (2-tailed)	.071	.642	.171	.830		.381
	N	54	54	54	54	54	54
Sophie Relatability	Pearson Correlation	.025	.182	.238	-.267	.122	1
	Sig. (2-tailed)	.856	.188	.083	.051	.381	
	N	54	54	54	54	54	54

Figure 4. *Correlation analysis of DLS scores with actor relatability.*

A correlation analysis of DLS scores, PHQ-9 scores and BDI-II scores was then performed.

Correlations

		DLS Score Sophie FB (seconds)	DLS Score Sophie TB (seconds)	DLS Score Hazel FB (seconds)	DLS Score Hazel TB (seconds)	PHQ_Score
DLS Score Sophie FB (seconds)	Pearson Correlation	1	.190	-.086	.076	.019
	Sig. (2-tailed)		.168	.536	.583	.892
	N	54	54	54	54	54
DLS Score Sophie TB (seconds)	Pearson Correlation	.190	1	.043	.128	.038
	Sig. (2-tailed)	.168		.755	.358	.788
	N	54	54	54	54	54
DLS Score Hazel FB (seconds)	Pearson Correlation	-.086	.043	1	.046	.058
	Sig. (2-tailed)	.536	.755		.741	.675
	N	54	54	54	54	54
DLS Score Hazel TB (seconds)	Pearson Correlation	.076	.128	.046	1	-.033
	Sig. (2-tailed)	.583	.358	.741		.811
	N	54	54	54	54	54
PHQ_Score	Pearson Correlation	.019	.038	.058	-.033	1
	Sig. (2-tailed)	.892	.788	.675	.811	
	N	54	54	54	54	54

Correlations

		DLS Score Sophie FB (seconds)	DLS Score Sophie TB (seconds)	DLS Score Hazel FB (seconds)	DLS Score Hazel TB (seconds)	BDI_Score
DLS Score Sophie FB (seconds)	Pearson Correlation	1	.190	-.086	.076	-.256
	Sig. (2-tailed)		.168	.536	.583	.062
	N	54	54	54	54	54
DLS Score Sophie TB (seconds)	Pearson Correlation	.190	1	.043	.128	-.108
	Sig. (2-tailed)	.168		.755	.358	.435
	N	54	54	54	54	54
DLS Score Hazel FB (seconds)	Pearson Correlation	-.086	.043	1	.046	.083
	Sig. (2-tailed)	.536	.755		.741	.551
	N	54	54	54	54	54
DLS Score Hazel TB (seconds)	Pearson Correlation	.076	.128	.046	1	-.156
	Sig. (2-tailed)	.583	.358	.741		.261
	N	54	54	54	54	54
BDI_Score	Pearson Correlation	-.256	-.108	.083	-.156	1
	Sig. (2-tailed)	.062	.435	.551	.261	
	N	54	54	54	54	54

Figure 5. *Correlation analysis of DLS scores, PHQ-9 scores and BDI-II scores.*

There was a negative correlation between the DLS score of Hazel's TB tasks and both BDI-II and PHQ-9 scores, this suggests that participants with higher PHQ-9 and BDI-II scores (more depressed participants) are more likely to consider Hazel an outgroup member as they cannot relate to her experiences. There was one unique finding from the analysis of BDI-II scores, finding a negative correlation between the DLS scores of Sophie and BDI-II score. This is quite unexpected as it suggests that people with a high BDI-II score (but not a high PHQ-9 score) consider other depressed persons as an outgroup.

A potential explanation for this is that the PHQ-9 is on a 27-point scale while the BDI-II is on a 63-point scale and assesses more depressive traits than the PHQ-9 including self-loathing (Kroenke, et al 2001) (Beck, et al 1996). Increased self-loathing in depression has been associated with loneliness (Ypsilanti et al, 2019) and Sophie referred to one of her interests as going out with friends. Potentially, participants with a high BDI-II score may have poorer iToM towards Sophie as she is depressed and has friends. Such a combination which may trigger high BDI-II scoring participants to feel more self-loathing as they have yet to overcome their depressive symptoms to make friends.

This study has found poorer performance on FB tasks, suggesting FB tasks require greater cognitive separation from reality (Döhnel et al, 2012). All conclusions drawn from correlation analyses must be considered tentatively as none reached

$p<0.05$ and as the adage goes, correlation does not equate to causation. That said, noteworthy trends have been drawn, such as the greater importance of relatability than the presence of depression in influencing iToM and intergroup bias. Based on this assumption we tentatively concluded that persons with more severe depression are likely to consider non-lonely, depressed persons as out-group members and have poorer iToM towards them due to this.

This study failed to find any potential cause of discrimination against depressed or mentally ill persons meaning further research must be performed into causes of this phenomenon.

What I have learned

Something important I learned is to just deal with problems. We had a lot of problems with the eye-tracker, the first this happened I spent more being angry about it than I should have. This only led to a lot of unneeded stress and didn't help, so I learned just to suck it up and deal with the problem. I also discovered the importance of controlling your working environment. The only desk space myself and other interns had access to was in the common room. Inevitably, I wasted a lot of time chatting. After a while, I decided to put my headphones on so I wouldn't be tempted to join a conversation- such a small thing which made a huge difference to my productivity. Another take away from this experience is to take the initiative. I ended up shadowing a master's student who was using the same paradigm as myself, so we shared participants and recording sessions. Initially we were going to work out a schedule to dictate who records which participant on what day, but this was quite inflexible. Instead, I suggested that we decide who will sit on which recording on a day-by-day basis. We tried this out and it worked well since it meant we could spend our time, our way and have more flexible working hours.

This internship ended up being a much bigger learning experience than I expected, I've taken many things from this experience big and small. It's increased my confidence significantly both in working and as a person.

Bibliography

Avenanti, A., Sirigu, A. and Aglioti, S.M. (2010). Racial Bias Reduces Empathic Sensorimotor Resonance with Other-Race Pain. *Current Biology*, [online] 20(11), pp.1018–1022. doi:https://doi.org/10.1016/j.cub.2010.03.071.

Beck, A. T., Steer, R. A,, & Brown, G. K. (1996). Manual for the Beck Depression Inventory-It. SanAntonio, TX: Psychological Corporation.

Brohan, E., Thornicroft, G., Rüsch, N., Lasalvia, A., Campbell, M.M., Yalçınkaya-Alkar, Ö., Lanfredi, M., Ochoa, S., Üçok, A., Tomás, C., Fadipe, B., Sebes, J., Fiorillo, A., Sampogna, G., Paula, C.S., Valverde, L., Schomerus, G., Klemm, P., Ouali, U. and Castelein, S. (2022). Measuring discrimination experienced by people with a mental illness: replication of the short-form DISCUS in six world regions. *Psychological Medicine*, 53(9), pp.1–11. doi:https://doi.org/10.1017/s0033291722000630.

Döhnel, K., Schuwerk, T., Meinhardt, J., Sodian, B., Hajak, G. and Sommer, M. (2012). Functional activity of the right temporo-parietal junction and of the medial prefrontal cortex associated with true and false belief reasoning. *NeuroImage*, 60(3), pp.1652–1661. doi:https://doi.org/10.1016/j.neuroimage.2012.01.073.

Gronholm, PC., Kline, S., Lamba, M., Lempp, H., Mahkmud, A., Cano, GM., Vashisht, K., San Juan, NV. and Sunkel, C. (2024). Exploring perspectives of stigma and discrimination among people with lived experience of mental health conditions: a co-produced qualitative study. *EClinicalMedicine*, pp.1-10. url:https://www.thelancet.com/journals/eclinm/article/PIIS2589-5370(24)00088-9/fulltext.

Henderson, C., Potts, L. and Robinson, E.J. (2019). Mental illness stigma after a decade of Time to Change England: inequalities as targets for further improvement. *European Journal of Public Health*, 30(3), pp.497–503. doi:https://doi.org/10.1093/eurpub/ckaa013.

IBM Corp. 2022. IBM SPSS Statistics for Windows, Version 29.0.0.0. Armonk, NY: IBM Corp.

Kroenke, K., Spitzer, R.L. and Williams, J.B.W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, [online] 16(9), pp.606–613. doi:https://doi.org/10.1046/j.1525-1497.2001.016009606.x.

Molenberghs, P. and Louis, W.R. (2018). Insights From fMRI Studies Into Ingroup Bias. *Frontiers in Psychology*, 9. doi:https://doi.org/10.3389/fpsyg.2018.01868.

Quesque, F., Apperly, I., Baillargeon, R., Baron-Cohen, S., Becchio, C., Bekkering, H., Bernstein, D., Bertoux, M., Bird, G., Bukowski, H., Burgmer, P., Carruthers, P., Catmur, C., Dziobek, I., Epley, N., Erle, T.M., Frith, C., Frith, U., Galang, C.M. and Gallese, V. (2024). Defining key concepts for mental state attribution. *Communications Psychology*, [online] 2(1), pp.1–5. doi: https://doi.org/10.3389/fpsyg.2018.01868.

Southgate, V., Senju, A. and Csibra, G. (2007). Action Anticipation Through Attribution of False Belief by 2-Year-Olds. *Psychological Science*, 18(7), pp.587–592. doi:https://doi.org/10.1111/j.1467-9280.2007.01944.x.

Systems , S. (n.d.). *Sona Systems : Cloud-based Subject Pool Software for Universities*. [online] www.sona-systems.com. Available at: https://www.sona-systems.com/.

Thornicroft, G., Rose, D., Kassam, A. and Sartorius, N. (2007). Stigma: ignorance, prejudice or discrimination? *British Journal of Psychiatry*, [online] 190(3), pp.192–193. doi:https://doi.org/10.1192/bjp.bp.106.025791.

Tobii AB (2021). Tobii Pro Lab (Version 1.171) [Computer software]. Danderyd, Sweden: Tobii AB.

Wilder, D.A. (1978). Reduction of intergroup discrimination through individuation of the out-group. *Journal of Personality and Social Psychology*, 36(12), pp.1361–1374. doi:https://doi.org/10.1037/0022-3514.36.12.1361.

Ypsilanti, A., Lazuras, L., Powell, P. and Overton, P. (2019). Self-disgust as a potential mechanism explaining the association between loneliness and depression. *Journal of Affective Disorders*, 243, pp.108–115. doi:https://doi.org/10.1016/j.jad.2018.09.056.