

Institut for Psykologi

STANDARDFORSIDE TIL EKSAMENSOPGAVER

Modultitel:	K08: Speciale	
Navn:	Niclas Ashkan Johansen	
Evt. samarbejdet med:		
Uddannelse: (BA. Sidefag eller KA.)	Kandidat.	
Termin: (f.eks. E19)	S24	
Vejleder:	Lotte Huniche	
Antal sider og typeenheder:	79.15 sider, 189.960 anslag	
Afleveret: (dato og klokkeslæt)		Modtagers underskrift:

Om brugen af kunstig intelligens som terapi

Kandidatspeciale i Psykologi, Syddansk Universitet

Af Niclas Ashkan Johansen

Vejleder: Lotte Huniche

Abstract

Introduction. More and more people seem to be using artificial intelligence such as ChatGPT, so called *large language models* (LLMs), as a form of therapy. Given the capabilities of these models, it seems to be a qualitatively different phenomenon from earlier therapy chatbots. However, there is a dearth of research, particularly qualitative, that investigates this topic. Therefore, this thesis investigates the experiences that reddit users have had using LLMs as therapy. **Methods.** The research questions is explored through five semi-structured interviews with five informants recruited from forums about *large language models* on reddit.com. The approach is an Interpretative Phenomenological Analysis, drawing on Husserl's conception of the I-and-thou relationship and Sartre's of theory of "the gaze", as well as a meta-theoretical framework on therapy, the common factors described by Bailey & Ogles. **Analysis and discussion.** Through analysis of the interviews, five themes and two subthemes are identified: 1) Changes, 2) Insight and relief, 3) AI: subject or tool?, 3.1) The role of communication and 3.2) A sort of personality, 4) Anxiety about judgement and 5) A central place in the lifeworld. The analysis sheds light on how the informants experience having used LLMs as therapy has changed their lives for the better in ways very similar to therapy, through insight and relief, but also the ambivalences they experience regarding the 'therapeutic' relationship.' This is interpreted through the application of Husserl's I-thou relationship and how it meets some criteria, but not all, for such a relationship. The anxiety about the judgement of others and how it is absent with LLMs, is interpreted through Sartre's idea of the "gaze," which gives a potential explanation for a) why the informants find it easier to confide in LLMs, but b) also less meaningful and significant. A risk of alienation and loneliness is discussed. Finally, the central place in the lifeworld is analyzed, which seems to entails both a feeling of safety but also less autonomy or self-efficacy, which reveals a fundamental difference between the goals of traditional therapy and LLMs used as therapy. **Conclusion:** LLMs used as therapy is similar to traditional therapy with regard to some of the changes, insights and relief, that it yields; on the other hand, it is different with regard to the relationship, and the significance and meaningfulness of the experience. Furthermore, it seems to entail risks of alienation and dependence. **Implication:** Further studies are proposed to better understand the topic: both a mixed methods approach to contrast and compare the relative importance of knowing that one is communicating with a human means to a client, as well as longitudinal studies to understand potential risks of alienation and dependence.

Indholdsfortegnelse

Introduktion.....	3
Begrundelse for emnet og forskningsspørgsmål	3
<i>Problemfeltet</i>	4
<i>Problemstilling</i>	6
<i>Forskningsspørgsmål</i>	7
Forskningsspørgsmålets led og relevant baggrundsviden hertil.....	8
<i>Hvilke oplevelser</i>	8
<i>Reddit-brugere</i>	8
Large language models såsom ChatGPT.....	8
Som terapi.....	11
<i>Bailey & Ogles fem faktorer</i>	13
Teori.....	16
<i>Om IPA generelt</i>	16
<i>Fænomenologi</i>	16
<i>Hermeneutik</i>	20
<i>Det ideografiske</i>	23
<i>Husserls begreb om jeg-og-du relation</i>	24
Metode	28
Om valget af kvalitativ metode	28
Om valget af IPA	29
Om valget af det semi-strukturerede interview	29
Om valget af kvalitetskriterier.....	30
<i>Elliot, Fischer og Rennies (1999) kvalitetsindikatorer</i>	31
<i>Tracys (2010) kvalitetskriterier</i>	32
Litteratursøgning	33
<i>Systematisk søgning</i>	33
Forskningsdesign.....	34
<i>Om udformningen af interviewguiden</i>	35
Om rekruttering	36
<i>Adgang til felten</i>	37
Rammesætning for interviewene.....	39
<i>Informeret samtykke og fortrolighed</i>	39
<i>Briefing</i>	40
<i>Konsekvenser af deltagelse</i>	40
<i>Debriefing</i>	41
Transskription.....	42
Andre etiske overvejelser	43
<i>Mikroetiske overvejelser</i>	43
<i>Makroetiske overvejelser</i>	45
Forforståelse og bias.....	46
Analyse og diskussion.....	49
Analysens trin i IPA.....	49
<i>Første trin: fordybelse i empirien</i>	49
<i>Andet trin: begyndende noter</i>	49
<i>Tredje trin: Udvikling af opståede temaer</i>	52
<i>Fjerde trin: undersøgelse af forbindelser imellem opståede temaer</i>	54
<i>Femte trin: bevægelsen videre til næste case</i>	54
<i>Sjette trin: at lede efter mønstre på tværs af cases</i>	55
Situering af informanter	55

Tabel over fremanalyserede temaer på gruppeniveau	56
Analyse og diskussion	59
<i>Overordnet tema 1: Forandringer</i>	60
<i>Overordnet tema 2: Indsigt og lindring</i>	62
<i>Overordnet tema 3: Den kunstige intelligens – subjekt eller redskab?</i>	69
<i>Overordnet tema 4 – Angsten for andres dom</i>	80
<i>Overordnet tema 5 – En central plads i livsverdenen</i>	88
Diskussion af metode	90
<i>Specificering af perspektiv</i>	90
<i>Kohærens</i>	90
<i>Situering</i>	91
<i>Sondring imellem generelle og specifikke formål</i>	92
<i>Eksempler</i>	93
<i>Troværdighedstjek</i>	93
<i>Etik</i>	93
<i>Et værdigt emne og et signifikant bidrag</i>	94
Konklusion og perspektivering	96
Konklusion	96
Perspektivering	98
Referenceliste	101
Bilag	112
1. Rekrutteringsopslag	112
2. Samtykkeerklæring	113
3. Interviewguide	115
4. Opståede temaer på individniveau	117
4a – Julia	117
4b - ROBERT	125
4C - IMRAN	127
4D - Anna	130
4E - Martin	134

Introduktion

Indeværende speciale har til formål at undersøge følgende: hvilke oplevelser som reddit-brugere, der har benyttet kunstig intelligens – mere specifikt *large language models* – såsom ChatGPT til terapi, har haft med dette. Dette bliver gjort igennem fem semi-strukturerede interviews med fem informanter rekrutteret via hjemmesiden reddit.com.

Det teoretiske afsæt for opgaven er tilgangen *Interpretative Phenomenological Analysis* (herefter benævnt IPA). I forlængelse af dette vil der blive benyttet fænomenologiske begreber og teorier fra særligt Husserl og Sartre, herunder og henholdsvis ”jeg-og-du-relationen” og ”blikket”, men også et meta-teoretisk framework om terapi, de fem fælles faktorer som Bailey & Ogles (2023) beskriver.

Dette – navnlig teori og metode - bygger videre på undersøgelser, som jeg først piloterede med et metodeprojekt (Johansen, 2024).

Begrundelse for emnet og forskningsspørgsmål

Det ville ikke være en overdrivelse, at sige, at OpenAI med lanceringen af ChatGPT til den brede offentlighed i 2022, har startet noget, der ligner en videnskabelig revolution (Bubeck et al, 2023; Ciesla, 2024).

Denne generative kunstige intelligens er i stand til at producere tekst, svar på spørgsmål, kode, digte, billeder, og et utal af andre ting, af en sådan kvalitet at de ikke oftest ikke er til at sondre imellem det og noget et menneske har produceret (Bubeck et al, 2023).

Turing-testen (Turing, 1950), som stipulerer at der er tale om kunstig intelligens, når en blændet observatør til en samtale imellem menneske og maskine, ikke længere pålideligt kan sondre imellem hvem der er menneske og hvem der er maskine, er for længst bestået (Bubeck et al, 2023; Ayers et al, 2023).

Og når det er til at skelne, så vil det ofte være fordi det, den kunstige intelligens producerer er af *højere* kvalitet, end det man ville forvente at se hos et menneske (Bubeck et al, 2023). Det er da heller ikke kun Turing-testen, der er bestået.

ChatGPT har foreløbig bl.a. bestået så forskellige eksaminer som en eksamen i jura ved University of Minnesota (Choi et al, 2023), den amerikanske medicinske embedseksamen United States Medical Licensing Examination (Kung et al, 2023), en eksamen på MBA-niveau ved Wharton (Terwiesch, 2023), samt et væld af andre prøver.

Antallet af brugere taler da også for sig selv: med over 100 millioner brugere på knapt tre måneder fra lanceringen, har den sat rekord som det den hurtigst voksende software applikation nogensinde (Hu, 2023).

En søgning på Scopus afspejler den samme virkelighed indenfor academia: siden offentliggørelsen af ChatGPT frem til tidspunktet hvor jeg skriver denne her sætning (16. marts), er der blevet publiceret 5.869 fagfællebedømte artikler og 2.711 preprints, hvor ordet ChatGPT indgår i titel, søgeord og/eller abstract.

Artiklerne begrænser sig ikke bare til computer science, men spænder over alt fra medicin, litteraturvidenskab, jura, molekylærbiologi og et hav af andre emner. Hvad mere er, denne aktivitet bliver kun yderligere accelereret af konkurrerende kunstige intelligenser, såsom Googles Gemini, Antropics Claude, samt et væld af andre. Selvsagt, så er det med den mængde aktivitet, svært at følge med i udviklingen, selv for eksperter indenfor feltet (What's next in the , 2023).

Problemfeltet

Ærindet med dette speciale, er at bidrage med beskrivelse og analyse af den påvirkning, som denne nye generation af kunstig intelligens har på psykologien, og mere specifikt terapi. Givet hvor nyt fænomenet er, har det ikke været muligt for mig at opdrive pålidelige tal fra fagfællebedømte publikationer, men at dømme ud fra nyhedsindslag, ser det ud til at flere og flere personer, bruger disse som terapi (Tidy, 2023).

Eksempelvis kan det anføres hvordan character.ai's chatbot 'Psychologist' har modtaget mere end 78 millioner beskeder, siden den blev skabt for lidt over et år siden, samt at der er mere end 475 chatbots med navne hvor ordene 'therapist', 'therapy', 'psychologist' og 'psychiatrist' indgår på samme hjemmeside (Tidy, 2023).

Ligeledes ses det ved et utal af kommentartråde på interfora som reddit.com, hvor brugere beretter om deres oplevelser med at anvende ChatGPT som terapeut og udveksler råd til *prompts*¹ (e.g. [Cairo-TenThirteen], 2023; [TurtleMaster472463], 2023). For mig at se er der to årsager til at dette fænomen er vigtigt for psykologer at forstå.

For det første, fordi der allerede nu er artikler, hvor der bliver spekuleret i om kunstige intelligenser à la ChatGPT en dag helt kan erstatte psykologer og psykiatere (e.g. Cheng et al, 2023). For det andet, ud fra en formodning – som også underbygges i empirien i indeværende speciale – om at nogle klienter enten vil supplere eller fravælge terapien med et menneske til fordel for en chatbot.

Hvad der motiverer dette supplement eller fravalg, samt hvilken slags oplevelser de har med det, er for mig at også relevant for klinikere forstå.

Der er en omfangsrig litteratur om brugen af chatbots i terapeutisk kapacitet, der strækker sig helt tilbage til 1960'ernes proto-chatbot ELIZA (Weizenbaum, 1966) og der har da også været lovende kliniske resultater i behandlingen af depression (Lim et al, 2022), substansbrugsforstyrrelser (Prochaska et al, 2021) og ADHD (Fitzpatrick et al, 2017) med chatbots.

Alligevel er det min overbevisning, at den nye generation af kunstig intelligens udgør et kvalitativt anderledes fænomen, end de chatbots der beskrives i selv forholdsvis nye artikler, og derfor er en risiko for en anakronistisk forståelse af fænomenet, hvis man forlader sig på denne litteratur.

¹ *Prompts* er de inputs, som brugerne giver til en generativ AI for at få en bestemt respons fra den.

Som eksempel på dette, kan anføres Haque og Rubya (2023), hvor den indsamlede empiri var fra perioden 1. januar 2019 og 1. maj 2022 – altså før ChatGPT blev offentligt tilgængelig – og de chatbots, der blev beskrevet, formodentlig benyttede en art beslutningstræer, snarere end de genererede helt originale svar som de nyere modeller (Graber-Stihl, 2023)

Som jeg vil forsøge at godtgøre med den empiri jeg har genereret og min tolkning af den, så mener dog der er en verden til forskel på disse to. Derfor har jeg i min oprindelige scoping søgning af emnefeltet (se afsnittet om litteratursøgning under metode), primært begrænset mig til artikler, der eksplicit beskæftiger sig med den nye generation af kunstig intelligens indenfor psykologi og terapi.

Overordnet kan de artikler inddeles i følgende kategorier:

a) Teoretiske artikler, såsom Sedlakova og Trachsels (2023), hvor de ud fra begrebsanalyser af eksempelvis agens, autonomi, bevidsthed og intentionalitet forsøger at redegøre for om kunstig intelligens overhovedet kan indgå i en terapeutisk relation, og konkluderer at svaret er nej.

b) Eksperimentelle paradigmer, hvor mennesker og kunstig intelligens sammenlignes i forhold til bestemte opgaver, såsom identifikation af selvmordsrisici i vignettestudier (Levkovich & Elyospeh, 2023), nøjagtighed og empati i svar på patientspørgsmål (Ayers et al, 2023), test af *emotional awareness* (Elyoseph et al, 2023), overensstemmelsen imellem adfærdsvurderinger af forumindlæg (Haman et al, 2023) eller vurdering af personlighedstræk (Rao et al, 2023).

3) Kvalitativ analyse af artefakter, såsom samtaler med ChatGPT, hvor forfatteren foregiver at være fx. klient eller en kliniker i vildrede i forhold til en behandlingsplan (Raile, 2024), for at vurdere den kliniske utilitet.

Problemstilling

Som jeg ser det er der dog en række begrænsninger med disse studiedesign, som jeg vil beskrive nedenfor:

1) Et indlysende problem med denne teoretisk afledte konklusion, er dog for mig at se at den strider med førstepersonsperspektivet, som ses i kommentartrådene anført ovenfor, hvor brugerne

netop selvidentificerer sig, som nogle der har brugt kunstig intelligens til terapi. Jeg tror, der er behov for studier, der ikke kun er teoretisk, men også empirisk funderet.

2) De empiriske studier har dog følgende begrænsninger. For det første, at selvom de undersøger egenskaber der er forbundet med og indgår i terapi – indlevelsesevne, kognitiv empati, *theory of mind*, etc. - så kan terapi heller ikke reduceres til disse (e.g. Wampold, 2015).

For det andet, så er der en begrænsning i at de kun tager udgangspunkt i tredjepersonsperspektivet: de vurderer hvor nøjagtigt fx en sindstilstand hos et menneske kan identificeres af kunstig intelligens, men ikke hvordan det føles for den givne person.

Det virker ikke selvindlysende sandt at det at ens følelser nøjagtigt bliver identificeret er det samme som at føle sig forstået – og derfor tror jeg der er behov for at forstå brugernes førstepersonsperspektiv (Johansen, 2024).

3) Udover de mulige indvendinger imod den økologisk validitet i den slags rollespil som Raile (2024) benytter, så er der det samme problem med ikke at få brugernes førstepersonsperspektiv, som under der er ved artiklerne under 1) og 2), idet analysen går kun på de sproglige artefakter, som ChatGPT genererer.

Med disse begrænsninger i mente, mener jeg altså at der er behov for en kvalitativ tilgang for at netop at få det førstepersonsperspektiv på hvordan terapi med kunstig intelligens opleves af brugerne selv - og at et interviewbaseret studie er den mest oplagte måde at få dette (Howitt, 2016).

Selv hvis man ikke køber min præmis, om at de nyere kunstige intelligenser er væsensforskellige fra tidligere chatbots, så påpeges det da i øvrigt i et review af sidstnævnte emne (Balcombe, 2023) at der mangler kvalitative, interviewbaserede undersøgelser af samme. Med andre ord, det forekommer mig at der uanset hvad er en lakune i vores viden om dette fænomen.

Min forhåbning er at dette speciale kan være med til at udfylde denne lakune.

Forskningsspørgsmål

I forlængelse af dette lyder mit forskningsspørgsmål således:

Hvilke oplevelser har redditbrugere, der har benyttet large language models som ChatGPT som terapi, haft med dette?

Forskningsspørgsmålets led og relevant baggrundsviden hertil

I det følgende vil jeg uddybe og begrunde forskningsspørgsmålets led, samt supplere med relevant baggrundsviden, hvor det forekommer nødvendigt.

Hvilke oplevelser

Hvad jeg sigter til med denne del af forskningsspørgsmålet er fortrinsvis to ting: for det første at ærindet med undersøgelsen er åbent og eksplorativt og ikke hypotesedrevet, for det andet at jeg er interesseret i førstepersonsperspektivet. Hvordan det knytter an til mit valg af IPA som metodisk og teoretisk standpunkt vil blive uddybet under afsnittet *Vidensskabsteori i IPA*.

Reddit-brugere

Reddit er en amerikansk hjemmeside, der med sin 430 millioner aktive månedlige brugere er en af de mest populære hjemmesider (Pew Research, 2019). Der er tale om en platform med tusindvis af brugerskabte fora (såkaldte 'subreddits') dedikeret til bestemte interesser, hvor brugerne kan oprette, kommentere og stemme på emner og indhold.

Demografisk er brugerne hovedsageligt mandlige (59 %), hvide (70 %), vestlige (61.4 % er amerikanere) og veluddannede (42 % har minimum en bachelor) (Perrin 2019). Eller med Henrich et als (2010) akronym: *WEIRD people* – *western, educated, industrialized, rich, democratic*. Når jeg i opgaven beskæftiger mig med denne gruppe og kun denne gruppe er årsagen enkel: det var her, det var muligt for mig at rekruttere informanter (se afsnittet om rekruttering under metode flere detaljer).

Large language models såsom ChatGPT

Selvom betegnelsen kunstig intelligens ofte bruge synonymt med *large language models* såsom ChatGPT, så er dette er dog kun en type af kunstig intelligens (Ciesla, 2024). Når

betegnelsen kunstig intelligens (eller synonymer som *artificial intelligence* eller *AI*), så refererer de dog i denne opgave altid til disse *large language models* (herefter LLMs).

Udgangspunktet for disse er enorme mængder tekst – i størrelsesordenen mange milliarder sider – hentet fra internettet, som fx bøger, indlæg på sociale medier, artikler, hjemmesider, etc, som bruges som træningsdata. Heraf kommer navnet *large language models* også (Ciesla, 2024).

Denne data brydes ned til såkaldte *tokens* – korte ord eller bestanddele af ord – som tildeles en given numerisk værdi. Igennem en træningsproces lærer transformerarkitekturen, en bestemt type af neurale netværk – en struktur der er analog til neuronerne i den menneskelige hjerne – at forudsige den sandsynlige sammenhæng imellem disse *tokens* (Radford et al, 2019).

Dette sker under såkaldt *unsupervised learning*, hvor der ikke er menneskelig involvering: modellen sammenholder det forudsagte *token* med det der faktisk findes i dens træningsdata; ved uoverensstemmelser imellem disse foretages der såkaldt *backpropagation* igennem netværket, hvor modellens parametre i det neurale netværk justeres. Modellen forbedres således iterativt (Radford et al, 2019; Ciesla, 2024).

Igennem denne proces lærer modellen iterativt, indtil den også kan generalisere de underliggende mønstre i sproget, sådan at de kan generere tekst til inputs der ikke er i dens træningsdata. Denne proces giver dog kun én en såkaldt *base model* – en slags dokumentgenerator – og er sig selv ikke tilstrækkeligt (Karpathy, 2023)

For at modellerne kan fungere som assistenter som ChatGPT gør, er det også nødvendigt med *fine tuning* (Ciesla, 2024). Dette gøres blandt andet ved at erstatte det oprindelige datasæt med et nyt datasæt, der er mindre, men kurateret; fx titusindvis af sider med både et spørgsmål og et idealtypisk svar på det givne spørgsmål, forfattet af forskellige eksperter indenfor et væld af forskellige områder (Ciesla, 2024; Karpathy, 2023).

Samtidig med at sammenhængene fra det oprindelige datasæt bevares, så lærer den nu også at imitere det svar og responsformat, der findes i det nye datasæt – den lærer således at svare

relevant på en brugers *prompts* (Ciesla, 2024). Dette suppleres yderligere med bl.a. *reinforcement learning from human feedback*, hvor menneskers bedømmelser om et svar er godt eller dårligt fører til yderligere finjustering af modellen (Ouyang et al, 2023).

Det er i kraft af disse trin, at modellerne er i stand til at gøre de ganske enkelt forbløffende ting, de kan, som fx at komponere rimede digte, der fører matematisk bevis for den uendelige mængde primtal, besvare spørgsmål der ville forudsætte avanceret *theory of mind* i mennesker eller programmere på et niveau, der kunne sammenlignes fordelagtigt med den gennemsnitlige softwareingeniør (Bubeck et al, 2023).

Hvordan modellerne 'tænker.'

I kraft af modellernes arkitektur – utallige lag af neurale netværk, deraf også navnet deep learning – er kun input og output er synligt for os, og dermed udgør modellerne en slags black box selv for deres udviklere (Mitchell & Krakauer, 2023). Dette gør det vanskeligt afgøre præcis hvad deres processering egentlig dækker over – om de egentlig 'tænker' eller 'forstår' på en ikke-triviell måde som anførselstegnene også antyder (Mitchell & Krakauer, 2023).

I en rundspørgen er eksperterne indenfor området delt omtrent på midten med hensyn til spørgsmålet om hvor vidt de nogensinde kunne gøre det ved blot at blive trænet på tekst (Mitchell & Krakauer, 2023).

For nuværende synes det dog, selvom det LLMs kan er fuldstændig forbløffende, rimeligt at betragte det som 'statistik på steroider.' De mangler for det første bevidsthed: de har ikke en subjektiv oplevelse af at være til, der er ikke noget det *er ligesom* at være en LLM, hvad der også udelukker intentionalitet, da det forudsætter en bevidsthed der kan være rettet mod noget (Butlin et al, 2023; Zahavi, 2018).

I tråd med dette, så er deres 'kognition' også forskellig i og med at de ikke er kropsliggjorte væsener som os, selv når de styrer robotsystemer (Mitchell & Krakauer, 2023), hvad der gør at de mangler emotioner, og heller ikke besidder koncepter eller repræsentationer af verden, der er funderet i sensoriske oplevelser (Butlin et al, 2023).

Det sidste ses bl.a. ved hvordan nogle rumlige opgaver, som ville være trivielle for mennesker, som fx hypotetisk at stable objekter ovenpå hinanden på en stabil måde, er forbløffende svært for fx ChatGPT (Bubeck et al, 2023), især sammenlignet med hvor godt den klarer sig på andre og for mennesker langt sværere opgaver.

Endelig mangler de også autonomi: hvor et menneske besidder evnen til at selvstyrende, fleksibel handling og læring, hvad der gør at vi kan holdes moralsk ansvarlige, så kan en LLM kun forfølge givne mål (Butlin et al, 2023).

Da disse egenskaber – bevidsthed, intentionalitet, kropsliggjorthed, autonomi – ofte figurerer som nøglebegreber i fænomenologien, har dette også bidraget til at IPA blev valgt som teoretisk ståsted; disse begreber vil blive udfoldet i større grad i afsnittet om *Videnskabsteori i IPA*.

Som terapi

Når jeg har brugt konjunktionen 'som' og ikke adverbiet 'til' i mit forskningsspørgsmål, så skyldes det et hensyn, som jeg tror måske vil blive tydelige for læseren, når præsenterer dem for et uddrag af APA's (2024) ordbogsdefinition på psykoterapi:

"Psychotherapy refers to any psychological service provided by a trained professional that primarily uses forms of communication and interaction to assess, diagnose, and treat dysfunctional emotional reactions, ways of thinking, and behavior patterns."

Nøgleordene er her "provided by a trained professional." Ligesom Sedlakova og Trachsels (2023) teoretiske analyse af nødvendige og tilstrækkelige betingelser, så udelukker det pr. definition og på forhånd at det skulle være muligt at der var tale om terapi imellem et menneske og en LLM – hvad der som sagt tilsyneladende strider med brugernes erfaringer.

Hensigten med ordet 'som' er altså at indledningsvist forholde mig agnostisk til dette spørgsmål, idet det kan forstås i betydningen 'som om', hvorimod 'til' ville fungere som en tilkendegivelse af at det var terapi².

Et andet og beslægtet problem med definitionen er dog dens vaghed: i og med at der tales om at det er "any psychological service," så ville et oplagt spørgsmål opfølgende spørgsmål være hvad præcis de består i?

At der bruges senere uddybes at disse psykologiske ydelser bruges i udredning, diagnosticering og behandling, fortæller os stadig ikke hvad de konkret består i og af. At der er et væld af meget forskellige terapeutiske retninger og skoler, gør selvfølgelig denne strategisk vaghed forståelig. Samtidig er den dog ikke særlig anvendelig til mit formål.

De fælles faktorer

Med dette problem i mente – altså at afgrænse og definere hvad der er og ikke er terapi – så tænker jeg, at man med fordel kunne benytte i litteraturen om de fælles faktorer i terapi. Udgangspunktet for disse, er den debat der har verseret i snart et århundrede (e.g. Rosenzweig, 1936): når terapi virker, hvad er det så, der gør det virksomt? Hvad er forandringsmekanismerne? (Wampold, 2015).

Hvor nogle tilskriver det meste af virkningen til specifikke komponenter såsom en given teknik (e.g. Hofman & Barlow, 2014), tilskriver andre den til det, som går igen på tværs af terapeutiske skoler og retninger – de såkaldte fælles faktorer (e.g. Wampold, 2015). I de fælles faktorerers favør er den kendsgerning, at meget forskellige terapeutiske retninger lader til at have slående ens udfald – den såkaldte 'dodo verdict' (Wampold, 2015; Bailey & Ogles, 2023)

Selvom der findes en betydelig mængde litteratur, der kritiserer de metodologiske begrænsninger ved studierne, der giver belæg for de fælles faktorer, så synes det værd at bemærke

² Selvom ordet 'terapi' kan dække over flere typer behandling (e.g. genterapi eller fysioterapi), så har jeg for nemheds skyld valgt at bruge det som synonym med psykoterapi og samtalerapi, hvad der både er den dagligsproglige brug, men også ofte ses i akademisk litteratur (e.g. Wampold, 2015; Bailey & Ogles, 2023).

følgende: selv de, der mest ihærdigt forsvarer betydningen af specifikke komponenter, medgiver samtidig også at de fælles faktorer er betydningsfulde (e.g. Hofman & Barlow, 2014).

Derudover er der også en pragmatisk årsag til at benytte litteraturen om de fælles faktorer, idet jeg får et meta-teoretisk framework, der gør sig gældende for alt *bona fide* psykologisk behandling, en slags liste med familieligheder, som jeg kan sammenholde med informanternes oplevelser.

Jeg tror dette faciliterer en mere nuanceret forståelse af fænomenet, hvor der både kan være overlap og forskelle, end en definition som enten er eller ikke mødt.

Derfor har jeg valgt at benytte de fem faktorer som beskrives af Bailey og Ogles (2023), som er en syntese af faktorer beskrevet af en lang række andre forskere, herunder Frank (1973), Goldfried (1980), Tschacher et al (2014) og Wampold (2015).

Bailey & Ogles fem faktorer

Ifølge Bailey & Ogles (2023) kendetegnes alt *bona fide* psykologisk behandling af følgende fem faktorer: 1) den terapeutiske relation, 2) motivation, 3) korrigerende erfaring, 4) indsigt og 5) *self-efficacy*. Nedenfor vil jeg dem.

1) Den terapeutiske relation og dens bestanddele, er oftest det, der pålideligt er fundet at være den mest virksomme del af terapi (Wampold, 2015), og skoler som den klientcentrerede ser den endda som det eneste nødvendige (e.g. Rogers, 1951). Den består først og fremmest af *båndet* imellem terapeut og klient, en tillidsfuld og varm relation, som også afhjælper social isolation, som er særligt betydningsfuld for klienter med kaotiske og forarmede relationer (Wampold, 2015).

Derudover består den af det det, der kaldes den terapeutiske alliance: enighed om mål og midler i terapien (Ogles & Bailey, 2023). Den terapeutiske relation forudsætter egenskaber som empati, varme, ydmyghed og oprigtighed fra terapeutens side (Ogles & Bailey, 2023). Den involverer ”evolved characteristics of humans as the ultimate social species; as such, psychotherapy is a special case of a social healing practice” (Wampold, 2015, s. 270).

2) Motivation er egentlig at regne for en klientfaktor, men den er samtidig uadskilleligt indviklet i den terapeutiske relation (Ogles & Bailey, 2023). Den dækker for det første over klientens ønske om bedring, men også deres ambivalens om forandring. Igennem accepten af den ambivalens kan terapeuten facilitere udforskning af den og hvad der taler for forandring -og dermed bistå med at øge klientens motivation for den, samt bistå i planlægningen af konkrete mål for at opnå den (Ogles & Bailey, 2023).

En anden vigtig bestanddel af denne faktor, er forventningseffekter. Klientens forventning om at terapien vil føre til bedring, er også med til at skabe en bedring på samme måde som placebo fungerer (Wampold, 2015). Men disse forventningseffekter er også modtage for interventioner fra terapeutens side: ved at give et troværdigt rationale for terapien, kan det både styrke klientens forventning om at terapien vil hjælpe og at de selv kan udføre de handlinger, der skal til. Sidsnævnte dækker over en øget følelse af *self-efficacy*, hvad der også kan blive selvopfyldende (Bailey & Ogles, 2023). Den er en meget robust prædikator for succes i terapi (Wampold, 2015).

3) Korrigerende erfaren dækker over klienten oplever sine emotioner på en ny måde, der dermed afkræfter deres udtalte eller udtalte forventninger og bidrager til *remoralisering* – et håb og forventning om bedring (Bailey & Ogles, 2023). Den er ofte vekselvirkende med indsigten, men alligevel også distinkt fra denne – den er følt og ikke kun tænkt, en oplevelse snarere end en forklaring (Bailey & Ogles).

Typisk vil denne – ligesom indsigten – blive faciliteret igennem teknikker såsom refleksiv lytning og åbne spørgsmål (Hill, 2020) Den involverer at *konfrontere*, at bevæge sig imod det ubehagelige snarere end at undgå det, som det ses ved fx eksponering eller desensitisering; at *udtrykke* vanskelige følelser og tanker, især dem der har været hæmmet og undertrykt; at regulere følelsernes intensitet, således at klienten også får en oplevelse af bemestring, der igen kan føre til *self-efficacy*. Konstruktet er ikke så grundigt undersøgt som den terapeutiske relation, men noget evidens peger imod emotionelle oplevelser som stærkere prædiktorer end selv sidstnævnte for udfaldet af terapi (Bailey & Ogles, 2023).

4) Indsigten er som nævnt beslægtet med den korrigerende erfaren, men er mere kognitiv end affektiv – den involverer en ny forståelse af ens problem (Bailey & Ogles, 2023). Det kan ske igennem en mere sokratisk *opdagelse*, hvor klienten selv bliver mere bevidst om sig selv og sin situation igennem fx selvmonitorering (Bailey & Ogles, 2023) og de førnævnte teknikker (Hill, 2020), men også igennem *læring* af ny information (fx som ved psykoedukation) (Bailey & Ogles, 2023).

Endelig kan *feedback* terapeuten giver, kompensere for den asymmetri, der er imellem hvad vi ved om os selv og hvad andre ved om os: klienten kan blive opmærksom på mønstre i deres følelser, tanker og handlinger og hvordan de indvirker på andre, som ellers ikke ville være synlige for dem (Vazire, 2010; Bailey & Ogles, 2023).

Selvom indsigt er konceptuelt til stede i alle terapier, så er det dog værd at bemærke at empirien kun synes at understøtte en moderat effekt af denne faktor som forandringsmekanisme (Bailey & Ogles, 2023).

5) Endelig er der *self-efficacy*, som der næppe findes nogen helt nøjagtig oversættelse for på dansk – begreber som selvtillid- og værd, autonomi, kompetence, etc., er kun delvist dækkende. Bandura (1997) definerede det således: [p]erceived self-efficacy refers to beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (s. 3).

Med andre ord: der er ikke en forvisning om at man nødvendigvis lykkes, men en forventning om at man kan udføre den handling der er nødvendig for at lykkes (Bailey & Ogles, 2023). Det er både en forandringsmekanisme i terapien, men også dens endelige mål: at klienten føler at de kan benytte affektive og kognitive indsigter i terapien til at løse deres problemer på egen hånd (Bailey & Ogles, 2023).

Som problemet med at oversætte betegnelsen *self-efficacy* muligvis illustrerede, så er det andet og mere end blot end følelse af kompetence: den indbefatter også en følelse af autonomi, at

selvom ens liv er formet af omstændigheder, så evner man også selv at forme omstændighederne – og at der med den frihed følger ansvar (Bailey & Ogles, 2023; Bandura, 1997). Dette gør også, at jeg ser nogle paralleller imellem terapiens mål og Sartres (1943) filosofi, som vil blive uddybet under næste afsnit om Teori i IPA

Teori

I de følgende afsnit vil jeg redegøre hvad der kendetegner IPA på et helt overordnet plan, for dernæst at gennemgå de teoretiske strømninger der trækkes på i IPA og dermed også dette speciale: fænomenologien, hermeneutikken og det ideografiske i den rækkefølge (Smith et al, 2009).

Derudover vil jeg også uddybe en række begreber fra særligt den fænomenologisk-eksistentielle strømning som benyttes i analysen: Husserls begreb om jeg-og-du relationen (Zahavi, 2023) og Sartres (1943, 1946) om frihed, ”blikket” i mødet med den anden og ond tro.

Om IPA generelt

IPA er en kvalitativ metode, der er karakteriseret ved og udviklet med henblik på, at forstå hvordan mennesker oplever fænomener (det fænomenologiske), men også hvordan de fortolker og tilskriver mening til disse (det fortolkende eller hermeneutiske) samt fællestræk og forskelle i hvordan de gør dette (det ideografiske eller særegne). Midlet er omhyggelige analyser af deres beretninger om deres levede oplevelser, typisk indsamlet via semi-strukturerede interviews (Smith et al, 2009).

Fænomenologi

Den første strømning, der kendetegner IPA er fænomenologien (Smith et al, 2009). Edmund Husserl betragtes oftest som grundlæggeren af fænomenologien, men den tæller også andre prominente figurer som Martin Heidegger, Jean-Paul Sartre, Maurice Merleau-Ponty og Emmanuel Lévinas, som hver især har bygget videre på det fundament, som Husserl grundlagde (Zahavi, 2018).

Der er altså ikke tale om et homogent filosofisk program, men samtidig er der også en række fællestræk hvad angår ideer og temaer (Zahavi, 2018). Det mest grundlæggende fællestræk, er opmærksomheden mod *hvordan* et givent fænomen fremtræder i bevidstheden, altså i vores førstepersons og subjektive oplevelse (Beyer, 2022).

Dette står i kontrast til fx et logisk positivistisk paradigme hvor man forsøger at afgøre *hvad* noget er 'objektivt' uafhængigt af en selv. I fænomenologien ville man afvise at dette overhovedet var muligt – alt videnskabelige erkendelse må også ske fra et førstepersonsperspektiv og bygge på det (Zahavi, 2018).

Dette er ikke en afvisning af objektivitet som sådan: pointen er at se på hvordan vores bevidsthed er gensidigt forbunden med verden, hvordan de gensidigt konstituerer hinanden, og at begge dele derfor må undersøges samtidig (Beyer, 2022; Zahavi, 2018).

Dette bringer os til Husserls ærinde, som til dels bliver opsummeret i mottoet 'til sagerne selv' (Zahavi, 2018). Ved omhyggeligt at undersøge oplevelser på dets egne præmisser, skal vi bevæge os væk fra en taget-for-given indstilling, vores dagligdags *livsverden* med dens forudindtagelser om og fortolkninger af fænomenet (Zahavi, 2018).

Dette må dog ikke forveksles med en rent deskriptiv tilgang til verden (Zahavi, 2018). At træde ud af det Husserl betegner den *naturlige attitude* er en refleksiv bevægelse, hvor vi bliver bevidste om vores bevidsthed og dens *intentionalitet* (Smith et al, 2009; Zahavi, 2018).

Intentionalitet dækker over hvordan bevidstheden altid er rettet mod *noget* og altid er det på en bestemt måde (Beyer, 2022). Måden kan være fremtrædelsesformen jeg intenderer noget – om jeg ser det direkte, drømmer om det, husker det, forestiller mig det, etc, men også hvordan jeg *intenderer* et givent objekts brug (Zahavi, 2018).

Dette må heller ikke forveksles med en snæver interesse for det rent subjektive: ideen med til at blive opmærksomme på vores eget intentionelle bidrag, til at fænomenerne fremtræder som de gør, er netop at forstå både det subjektive og det objektive, både erfaringsakten- og genstanden og sammenhængen imellem dem (Beyer, 2022; Zahavi, 2018)

Målet er at nå frem til de essentielle og invariante egenskaber ved fænomenet, som derved kan overskride det partikulære – og dermed også gøre sig gældende for andres oplevelse af fænomenet (Smith et al, 2009).

For at nå frem til dette, udviklede Husserl også en række metodologiske trin, som han anså for nødvendige for enhver fænomenologisk analyse (Smith et al, 2009; Zahavi, 2018). Først og fremmest igennem epoché, en sætten-i-parentes eller suspension (Smith et al, 2009).

Hvad der skal sættes i parentes er ikke bare vores forforståelser, vanetænkning og teorier, men selve naturlige indstilling om at den verden, vi erfarer, eksisterer afhængigt af os og det perspektiv vi har på den (Zahavi, 2018).

Derudover måtte man benytte *eidetisk reduktion og variation*: at sammenholde forskellige instanser af et givent fænomen og finde fællestrækkene, men også imaginært at forsøge at trække fra eller føje til, for at finde ud af hvor fænomenets grænser går og hvad de essentielle træk er (Smith et al, 2009).

Heidegger bidrog til fænomenologien både med et øget fokus på det eksistentielle og hermeneutikken, som gør den mere fortolkende end deskriptiv (Smith et al, 2009). Førstnævnte vil blive behandlet i indeværende afsnit, mens bidraget til hermeneutikken vil blive behandlet i afsnittet af samme navn.

Blandt de vigtigste begreber som Heidegger introducerede til fænomenologien er *dasein*, der ofte oversættes tilstedeværen, men i en bogstavelig oversættelse betyder 'der-væren' (Zahavi, 2018). Hvad det peger henimod er vores væren-i-verden og verdens væren i os: sondringen imellem det indre og ydre er uanvendelig, vores væren kan ikke forstås i et vakuum (Zahavi, 2018).

Det er netop i den givne kontekst vi er – vores medmennesker, objekterne, sprog og kultur – at bestemte handlinger bliver mulige og meningsfulde, og netop fokus på den enkeltes specifikke kontekst er da også af betydning for ideografien (Smith et al, 2009).

Merleau-Ponty betoner særligt i sin fænomenologi betydningen af kroppen og det at være kropsliggjort (Toadvine, 2023; Smith et al, 2009). I modsætning til den cartesianske dualisme der

hævder en opdeling imellem sjæl og legeme, at vi *har* kroppe, forfægter Merleau-Ponty denne opdeling og pointerer at vi *er* kroppe (Toadvine, 2023).

Det betyder at vores perception aldrig blot er passiv modtagelse af data, men altid kommer af et kropsligt engagement med verden, fx hvordan vi bevæger os eller vores øjne i forhold til det vi ser (Zahavi, 2018) I forlængelse af dette beskriver han også hvordan vores kognition er uadskillelig fra vores kroppe, hvad der er en del af grundlaget for feltet *embodied cognition* (Toadvine, 2023).

Sartre uddybede den eksistentielle dimension i fænomenologien som blev introduceret af Heidegger (Smith et al, 2009) og for indeværende speciale er der særligt to bidrag fra Sartre, der er af betydning. Hans analyser af den menneskelige friheds betydning i *Eksistentialisme er en humanisme* (Sartre, 1946) og hvad der sker med os i mødet med den anden i *Væren og Intet* (Sartre, 1943).

Med hensyn til friheden, er den i hans optik et grundvilkår for mennesket. Ifølge Sartre går vores eksistens forud for vores essens: det er når vi træffer et bestemt valg, at vi afgør hvilken slags mennesker vi er, ikke omvendt hvilken slags mennesker vi er, der afgør hvilke valg vi træffer (Sartre, 1946).

Ikke bare er vi frie, vi er dømt til at være frie (Sartre, 1946). Siden ingen Gud eller prædetermineret essens afgør hvilke valg vi træffer, så er ansvaret for vores valg alene vores ifølge Sartre (1946). Som ved Kants kategoriske imperativ, så må vi ifølge Sartre (1946) vælge sådan, at det kunne være et forbillede til efterlevelse for hele menneskeheden.

Selvom *fakticiteten* – omstændigheder som vores historiske og økonomiske kontekst, fortidens uforanderlighed, vores fysiske krop, etc. - kan begrænse hvilke valg vi kan træffe, så kan det aldrig udelukke at vi kan vælge (Sartre, 1946; Reynolds & Renaudie, 2022).

Dette ansvar er for Sartre forbundet med angsten, og vi kan blive fristede til at fornægte vores egen frihed – det selvbedrag han kalder ond tro (*mauvaise foi*) – med henvisning til *fakticiteten* (Sartre, 1943; Sartre, 1946; Atkins, 2008).

Disse begreber er også i spil i hans beskrivelse af mødet med den anden blik i *Væren og Intet* (Sartre, 1943). Hvad Sartre beskriver her, er hvad der sker når vi mærker den andens blik på os: det er netop ved 'blikket' at vi bliver selvbevidste idet vi antager den andens perspektiv på os og hvordan de dømmer os (Sartre, 1943).

"Blikket" både være en kilde til selvindsigt, men også til angst (Reynolds & Renaudie, 2022). Angsten hidrører først og fremmest fra påmindelsen om vores egen frihed, men også hvordan vores valg påvirker andre og hvordan vi står til ansvar for deres dom (Sartre, 1953; Reynolds & Renaudie, 2022).

Men angsten er også forbundet med den andens frihed og fremmedhed: vi kan ikke bestemme – både i betydningen "at være sikre på hvad den er" eller betydningen "beslutte den" – hvad den andens doms er (Sartre, 1943). Det indebærer en subjektiv oplevelse af tabt kontrol, som samtidig med at det kan være årsagen til en etisk opvågning, også kan gøre os uautentisk og konformt for at undgå den andens dom. Med andre ord, at vi lever i ond tro (Atkins, 2005).

Sartres opfattelse af mødet med den anden som grundlæggende konfliktfyldt, er da også blevet kritiseret af blandt andet Beauvoir og Merleau-Ponty (Atkins, 2005); når det er inkluderet i indeværende speciales analyse, er det ikke en utvetydig tilslutning til at mødet med den anden nødvendigvis *må* være sådan, men fordi det er en nyttig analyse til at forstå *når* mødet er sådan.

Hermeneutik

Den næste betydningsfulde strømning i IPA er hermeneutikken, hvoraf I'et for *interpretation* i akronymet stammer fra (Smith et al, 2009). Det er en teori om fortolkning, som oprindeligt blev appliceret i læsningen af bibelske tekster, men sidenhen er blevet anvendt indenfor et væld af andre discipliner (George, 2021), herunder IPA (Smith et al, 2009).

Hermeneutikkens kerne er fortolkningens metoder og formål, om og hvordan det er muligt at afdække en forfatters oprindelige hensigt på tværs af kontekst, og forholdet imellem den oprindelige kontekst og fortolkningens kontekst (George, 2021; Smith et al, 2009).

Nedenfor vil jeg skitsere tre af betydningsfulde teoretikere indenfor hermeneutikken: Schleiermacher, Heidegger og Gadamer, samt deres bidrag til teorien og relevansen for IPA specifikt (Smith et al, 2009).

Schleiermacher var blandt de første, der beskrev hermeneutikken systematisk. Han inddelte fortolkningen i den *grammatiske* og *psykologiske* (Smith et al, 2009).

Det grammatiske var om tekstens objektive betydning, samt hvordan forventninger og konventioner i forfatterens lingvistiske fællesskab gav den form og betydning; det psykologiske var om forfatterens egne hensigter og teknik, hvoraf tekstens særlige betydning også hidrørte (Smith et al, 2009).

I modsætning til senere post-moderne litteratur er der altså fokus *både* på det diskursive og det individuelle (Smith et al, 2009 som i IPA. Ligeledes – også i tråd med IPA – kunne den fortolkning ikke blot være et spørgsmål om mekanisk at følge regler, men involverede også håndværk og intuition (Smith et al, 2009).

Målet for denne fortolkning, var at forstå forfatteren bedre end vedkommende forstod sig selv (Smith et al, 2009). Selvom man indenfor IPA ville være overordentlig skeptisk overfor det sidstnævnte, så er det fuldt ud foreneligt med ideen om at den finkornede analyse i IPA bidrager med noget andet og mere end informanten selv nødvendigvis er bevidst om eller eksplicit formulerer (Smith et al, 2009).

Heideggers bidrag er i denne sammenhæng tofoldigt. For det første, at applicere hermeneutikken eksistentielt, på det levede liv, i stedet for blot indenfor kulturens produkter (George, 2021).

For det andet, at påpege fortolkerens egen kontekst spiller ind i den hermeneutiske proces (George, 2021; Smith et al, 2009). En fortolkning aldrig kan være antagelsesløs, som en udbredt tolkning af Husserls metode med epoché ellers forudsætter (Zahavi, 2018; Smith et al, 2009).

Vi bringer altid vores forforståelse ind i mødet med teksten, og ofte ville denne skjule egenskaber ved teksten for (Smith et al, 2009). Samtidig er det ofte i mødet med selve teksten at ens forforståelse bliver synlig for én som sådan – og dermed bliver fortolkningen tolkningen en refleksiv og cyklisk bevægelse imellem tekst og forforståelse (Smith et al, 2009).

I forlængelse af denne fortolkende tradition, understreger Gadamer netop tradition og histories betydning for fortolkningen, samt vekselvirkningen imellem fortolker og det fortolkede (Smith et al, 2009)

Som hos Heidegger, så synliggør engagementet med teksten den forforståelse man har; Gadamer nuancerer dette yderligere, ved også at påpege hvordan læsningen af teksten ikke bare synliggør men også har potentialet til at ændre forforståelsen og dermed fortolkningen (Smith et al, 2009).

Udover sit metodiske bidrag, er der også, for IPA, et værdifuldt forsvar for bestemte epistemiske normer og værdier (George, 2021). Som med den hermeneutiske cirkel, hvor del forstås ud fra helhed og omvendt (fx ord og sætning), så er fortolkende proces for Gadamer *cirkulært* (George, 2021).

Dermed er den fortolkende proces aldrig definitivt afsluttet. Vi vil altid være påvirket mere af tradition og fordom, end vi kan blive selvbevidste om. Vores forståelse kan altid blive dybere og dybere, og derfor må vi også være have en 'delfisk' bevidsthed om dens begrænsninger (George, 2021; Smith et al, 2009).

Hvordan Gadamer betoner forståelse frem for viden er også af betydning for IPA: i modsætning til det traditionelle ideal om replicerbar og akkumulativ viden indenfor den nomotetiske videnskab, så forsvares forståelsen som stammer fra fortolkningen, som en mere oprindelig form for sandhed, som vi til dels er blevet fremmedgjorte fra gennem den moderne videnskabs succes (George, 2021).

Dette knytter an til forskellen til forskellen mellem det ideografiske og det nomotetiske, som vil blive behandlet under næste afsnit.

Det ideografiske

Hvor den nomotetiske videnskab forsøger at opstille generelle og almentgyldige lovmæssigheder, der gælder på tværs af tilfælde og tid, så er det ideografiske en undersøgelse af det *partikulære* (Smith et al, 2009).

I modsætning til det meste moderne psykologi, hvor statistiske analyser på populationsniveau er *de rigueur*, så er der i IPA en forpligtelse på det ideografiske; denne forpligtelse på det ideografiske viser sig på i hvert fald to niveauer (Smith et al, 2009).

For det første, som i fokus på det partikulære i detaljen og hvor finkornet og systematisk analysen er; for det andet, i at man beskæftiger sig med partikulære oplevelser og fortolkninger hos enkelte personer (Smith et al, 2009).

Denne fremgangsmåde kan fungere eksplorativt og hypotese genererende, men er ikke begrænset til dette og kun dette (Flyvbjerg, 2010). At arbejde ideografisk indebærer ikke at man forkaster muligheden for at generalisere ud fra sin empiri (Flyvbjerg, 2010), men dog en mere forsigtig måde at nærme sig denne generalisering (Smith et al, 2009).

Oplagt kan en eller flere cases også fungere som ”sorte svaner”: selvom casen der ikke kan udledes prævalens ud fra en case, så kan den demonstrere eksistens – og dermed tjene til at falsificere en given teori (Flyvbjerg, 2010).

Men generaliseringen kan også ske triangulering af informanternes oplevelse og fortolkning af et givent fænomen (Howitt, 2016). Hvis informanterne, som alle er situerede og kropsliggjorte på forskellige måder, stadig har ens oplevelser af fænomenet, så giver det os en tentativ forståelse af fænomenet qua fænomen (Howitt, 2016).

Metoden hertil vil være en analytisk induktiv, hvor ens initielle hypotese afledt af en given case, sammenholdes med en ny case, og i lyset af denne revideres hypotesen så, indtil der foreligger en teori, der kan redegøre for alle eller de fleste af ens cases (Smith et al, 2009).

Husserls begreb om jeg-og-du relation

I det følgende vil jeg først beskrive hvad jeg-og-du relationen ifølge Husserl består af, og derefter vil jeg godtgøre hvorfor jeg mener at det forsyner en med et nyttigt begrebsapparat i forhold til at forstå terapi, især med henvisning til Wampold (2015), samt brugen af *large language models* til samme.

Hvad jeg-og-du relationen er.

Et godt udgangspunkt for at forstå Husserls begreb om jeg-og-du relationen, er at forstå empatiopfattelsen der lægger til grund for den – og også hvordan den er væsensforskellig fra *theory of mind* teorien, en udbredt redegørelse for empati indenfor kognitionspsykologien (Zahavi, 2023).

Theory of mind teorien hævder at empati foregår ved en slutning: ser jeg fx nogle græde, så slutter jeg mig til, at de er triste – og hvis jeg har en oplevelse af at det er noget andet og mere, så er det blot fordi slutningerne er blevet så automatiske, at de er blevet usynlige for mig som sådanne (Zahavi, 2023).

Ifølge Husserl derimod er ens forståelse af en andens sindstilstand ganske anderledes: den er umiddelbar, den er kropsliggjort og direkte (Zahavi, 2023). Dette udelukker selvfølgelig ikke at der er en kognitiv og slutningsbaseret andenforståelse – blot at den er sekundær til den førnævnte (Zahavi, 2008).

Denne opfattelse, der minder meget om den teoretiske retning om social perception, synes at stemme overens med både med *common sense* oplevelser i livsverdenen af at direkte erfare en andens affekt, men også evidens fra neurovidenskaben om spejlneuroner (Bonini et al, 2024; Cahoone, 2013).

I forbindelse med dette, er der også en anden vigtig forskel: hvor *theory of mind* teorien kan forstås som en slags *tilskuerteori* eller tredjepersonsperspektiv, så er der i Husserls empatiforståelse også mulighed for et andenpersonsperspektiv, et engageret perspektiv, hvor vi forholder os til en person som et *du*. (Zahavi, 2023).

Husserl opstiller en hel taksonomi over typer af empati, men jeg vil i det følgende nøjes med at beskrive de to, der virker mest vedkommende for indeværende, den unidirektionelle og den gensidige (Zahavi, 2023).

Den unidirektionelle involverer at A iagttager B og erfarer at vedkommende er trist, uden at B erfarer at A gør det. Den gensidige empati derimod, involverer at A er opmærksom mod B, B er opmærksom mod A, og at begge er opmærksomme på at den anden også er det (Zahavi, 2023).

I modsætning til de fleste *theory of mind* redegørelser for empati, som implicit ville betragte os som en slags observatører, der aflæser den andens sindstilstand uden at denne aflæsning påvirker vedkommende, så er der altså med Husserls *gensidige empati* muligheden for at A og B er i *berøring* med hinanden som det sker i et ansigt-til-ansigt møde (Zahavi, 2023).

Denne berøring kalder Husserl væren-for-hinanden, og den er grundlaget for at en jeg-og-du relation kan opstå. (Zahavi, 2023). Men en anden forudsætning er kommunikation, og denne kommunikation indebærer at sker en henvendelse fra den ene part til den anden, og at denne henvendelse tages op – henvendelsen alene er ikke nok (Eilan, 2024; Meindl & Zahavi, 2023).

Henvendelsen – fx en ordre, et løfte, en anmodning – må forstås som sociale handlinger. De er kun meningsfulde for så vidt at den anden tager henvendelsen op (Zahavi, 2023). Hvis og kun hvis dette sker, vil de to subjekter indgå i en slags fællesskab – Husserl pointerer da netop også de fælles etymologiske rødder som *community* og *communication* deler (Zahavi, 2023).

Her vil de to subjekter være kendetegnet af væren-i-hinanden i stedet for blot væren-for-hinanden: de to subjekter motiverer hinanden intentionelt, de ko-determinerer deres intentionalitet (Zahavi, 2023; Meindl & Zahavi, 2023).

I modsætning til fx Martin Bubers forståelse af jeg-og-du relationen, så er det vigtigt at pointere at Husserls forståelse af samme ikke nødvendigvis involverer en uhierarkisk og uinstrumentel indstilling til den anden (Zahavi, 2023). Som et minimum vil den dog have en anerkendelse af den andens autonomi, i og med at der i henvendelsen også må indebære at den anden kan afvise den eller svare igen (Zahavi, 2023).

Hvorfor det er vigtigt for indeværende.

For Husserl er der i hvert fald tre måder hvormed jeg-du relationen er af betydning: først og fremmest, for at forstå hvordan det objektive er intersubjektivt konstitueret, for det andet, for at forstå socialitet på både mindre og større skala, og for det tredje, i hvordan det er med til at skabe og forme selvet (Meindl & Zahavi, 2023).

Som Husserl beskriver det, så er det netop i mødet med et du at jeg bliver til et jeg, som er andet og mere end ren ego, mere end blot et subjekt med eksperientiel erfaring (Meindl & Zahavi, 2023; Zahavi, 2020). I mødet med den med den anden er jeg ikke bare bevidst om den anden, men også deres bevidsthed om mig – og dermed bliver jeg selvbevidst (Zahavi, 2015).

Det er netop i dyader af den type, at betingelserne skabes for egenskaber som sprog, repræsentation, værdier, kort sagt, mere sofistikeret kognition (Cahoone, 2013; Tomasello, 2020; Zahavi, 2020). Denne opfattelse af jeg-og-du relationens vigtighed flugter da også godt med evidens fra et væld af forskellige discipliner, såsom udviklingspsykologi, komparativ, evolutionær og klinisk psykologi (Schilbach et al, 2013; Cahoone, 2013; Tomasello, 2020; Hrdy, 2009).

Det er ikke muligt fyldestgørende at redegøre for denne litteratur i indeværende, men det billede, der tegner sig er et af hvordan delt opmærksomhed, delt intentionalitet og evnen til at percipere sindstilstande i artsfæller både fylogenetisk og ontogenetisk går forud for og udgør betingelsen for kompleks kognition og socialitet i mennesker, både som art og individ (Cahoone, 2013; Tomasello, 2020).

Netop dette knytter også an til Wampolds (2015, s. 270) ide om hvad terapi er og hvordan dens mekanismer involverer ”evolved characteristics of humans as the ultimate social species; as such, psychotherapy is a special case of a social healing practice.”

Af de fælles faktorer som Wampold (2015) tilskriver størst betydning, er da netop også empati og den terapeutiske alliance. På samme måde som Husserl ser empatien som grundlaget for jeg-og-du relationen, så anskues de også i litteraturen om de fælles faktorer for at være en af de vigtigste betingelser for at terapi overhovedet kan finde sted (Zahavi, 2023; Wampold, 2015).

På samme måde er den faktor der tilskrives suverænt størst betydning, da også den terapeutiske alliance: et bånd, der er kendetegnet samarbejdende, målrettet arbejde hvor der er enighed om mål og midler (Wampold, 2015) som synes at være et klokkeklart eksempel på den delte intentionalitet der finder sted i en jeg-og-du relation.

Med de sammenfald i mente, ville et oplagt spørgsmål her så være, hvad tilføjer begrebet om jeg-og-du relationen så, som et analytisk framework, som ikke findes i Wampolds model? Som jeg ser det supplerer den med en mere finkornet analyse af det menneskelige møde generelt. At det er generelt har den fordel, at det ikke er alle mine informanter, der har erfaring med traditionel terapi – og at deres sammenligning ofte vil være med 'den sociale helingspraksis' med personer, der ikke er skolede terapeuter.

Derudover, så opstilles der af Husserl også en række minimumsbetingelser for at dette møde kan finde sted: minimum to subjekter, som må have autonomi og intentionalitet, en gensidig og kropsliggjort empati, kommunikation hvor der er en henvendelse der bliver taget op, men som potentielt kunne afvises (Zahavi, 2023).

Da der, som det fremgår af afsnittet om LLMs er en forskel mellem dem og mennesker hvad angår subjektstatus (i.e. at de mangler subjektive mentaltilstande), autonomi, intentionalitet eller kropsliggjorthed, så tror jeg også at dette vil udgøre et frugtbart udgangspunkt for at forstå forskellene i informanternes oplevelse af LLMs vis-a-vis oplevelser med mennesker.

Metode

I dette afsnit, vil jeg gennemgå undersøgelsens metode og design, for at sikre troværdighed og transparens omkring de valg, jeg har truffet og rationalerne for hvorfor jeg har truffet dem (Elliot et al, 1999; Tracy, 2010). Før jeg beskriver design og metode mere i detaljen, vil jeg dog først begrunde valget af kvalitativ metode og IPA mere overordnet.

Om valget af kvalitativ metode

Som det bliver godtgjort under Litteratursøgning, er der tilsyneladende endnu ikke nogen interviewbaserede studier af brugen af de nye LLMs som terapi; det samme gør sig også gældende for de ældre generationer af chatbots til terapi (Balcombe, 2023). Et sådan studie kunne for mig at se tjene i hvert fald tre funktioner:

1. Hypotese generering: selvom casestudier og kvalitative studiers nytte ikke kun er begrænset til hypotese generering, så er det *også* en funktion, som de er formålstjenestelige til (Flyvbjerg, 2010). Givet hvor nyt fænomenet er, så synes jeg der er en klar utilitet i et forskningsdesign, der fordrer hypotese generering frem for hypotesetestning (Howitt, 2016). Med andre ord, så mener jeg at et kvalitativt studie kan være med til at kvalificere hvilke spørgsmål der bør stilles.

2. Førstepersonsperspektivet: i forlængelse af sidste sætning om oplevelserne mennesker har, så er en begrænsning ved de kvantitative studier angivet i introduktionsafsnittet, at

a) de kun beskæftiger sig med enkelte variable (e.g. emotional awareness) som terapi ikke kan reduceres til og

b) at de i kraft af deres forskningsdesign kun kan anvende tredjepersonsperspektivet, d.v.s. at de beskæftiger sig med hvor nøjagtigt LLM'erne identificerer fx følelser, men ikke hvordan det rent faktisk *opleves* af menneskerne (Johansen, 2024).

Ved at benytte kvalitativ metode og semi-strukturerede interviews, kan jeg derimod få derimod adgang til rig og dyb data individets oplevelse (Howitt, 2016) af et sammensatte fænomen.

Om valget af IPA

Netop IPA stemmer for mig at se overens med disse mål. I forhold til det hypotese genererende, så er der en række metodiske greb (fx epoché) i fænomenologien, der kan afstedkomme en mere antagelsesløs og mindre teoriladet tilgang, og i hermeneutikken en måde at forhøre og udfordre de forudforståelser, man uundgåeligt også har (Smith et al, 2009). Det ideografiske sigte bidrager derudover til det eksplorative sigte, ved at der også er øje for forskelle og ikke kun ligheder i informanternes beretninger (Smith et al, 2009).

I forlængelse af den tredje pointe i afsnittet foroven, så tager fænomenologien også førstepersonsperspektivet alvorligt, idet den ser det som udgangspunktet for alt viden (Zahavi, 2018). Ydermere har den også et nyttigt begrebsapparat for og analyser af meget af det jeg vil undersøge.

Jeg tænker her særligt på mødet med den anden i form af Sartres (1953) ”blikket” og Husserls ”jeg-og-du-relation” (Zahavi, 2023) som kan belyse hvordan det at henvende sig til en LLM er anderledes, samt beskrivelser af bevidsthed, intentionalitet, kropsliggjorthed, umiddelbar empati og autonomi – som netop er egenskaber, som LLMs mangler (Butlin et al, 2023).

Om valget af det semi-strukturerede interview

I forlængelse af mit ærinde om at få førstepersonsperspektivet på det fænomen, jeg var interesseret i, måtte jeg finde en metode der egnede sig til dette, en metode ville give mig rig og dyb empiri, jeg kunne analysere (Howitt, 2016).

Typisk vil dette i IPA blive opnået ved hjælp af dybdegående semi-strukturerede interviews, efterfulgt af informantens skriftlige beskrivelser af et givent fænomen, da disse giver adgang til førstepersonsperspektivet på en måde som fx observation principielt ikke kan (Howitt, 2016; Johansen, 2024).

Skriftlige beskrivelser blev da også overvejet som metode, da min rekruttering foregik på onlinefora, hvor mange brugere netop beskrev deres oplevelser med AI terapi. Jeg forkastede dog denne mulighed af følgende to årsager.

For det første, fordi beskrivelserne var for korte til at være formålstjenestelige (sjældent længere end tre-fire sætninger) og for det andet, fordi beskrivelserne de indeholdt, stort udelukkende forholdt sig til valensen af oplevelsen, i.e. om den føltes god eller dårlig, men ikke *hvorfor* og *hvordan* den var det.

For at generere empiri om netop dette, hvordan og hvorfor informanterne oplevede terapi med kunstig intelligens på en given måde og hvilken mening de tilskrev den, valgte jeg altså det semi-strukturerede interview som fremgangsmåde.

Udover førnævnte hensyn, var der også følgende rationaler for dette valg:

1) Det gav mig fleksibilitet, både i forhold til a) at forfølge spor, jeg ikke på forhånd vidste kunne være interessante, b) og i tråd med dette, at få uddybet og afklaret oplevelser som var selvfølgelig for informanten men nye og/eller fremmedartede for mig og c) at afvige fra interviewguiden, når de forudforståelser der lå til grund for mine spørgsmål (til trods for bestræbelser på at sætte disse i parentes) ikke stemte overens med informantens oplevelser.

2) I forhold til mine opstillede kvalitetskriterier, så tillod det mig også at lave troværdighedstjek i selve interviewet, hvor jeg kunne bede min informant forholde sig til en given tolkning og forståelse, jeg måtte have af deres svar (Brinkmann & Tanggaard, 2015).

Om valget af kvalitetskriterier

Som påpeget af Brinkmann og Tanggaard (2015), så må der, i og med at kvalitativ forskning har andre midler, mål og ambitioner end kvantitativ forskning, også nødvendigvis gælde andre kriterier for hvad kvalitet er i disse studier. Disse kriterier må respektere den kvalitative forsknings egenart (Thagaard, 2004).

Der er derimod ingen konsensus om *hvad* disse kvalitetskriterier bør være, og nogle synes endda gensidigt at udelukke hinanden (Tanggaard og Brinkmann, 2015): hvor nogle advokerer for endnu mere radikale brud med hele den logisk-positivistiske tradition, advokerer andre for at finde at finde ekvivalente kriterier til reliabilitet, validitet og generalisering (Thagaard, 2004).

Jeg har valgt at tage udgangspunkt i et udvalg af kriterier fra Elliot, Fischer og Rennie (1999) og Tracy (2010). Nedenfor følger en kortfattet beskrivelse af dem.

Elliot, Fischer og Rennies (1999) kvalitetsindikatorer

1. Specificering af perspektiv: her sigtes til at der bør være transparens for læseren om hvad det (videnskabs)teoretiske ståsted er, personlige forforståelser, metodiske begrundelser og målet med studiet (Elliot, Fischer og Rennie, 1999).

2. Situering af deltagere: at der indgår beskrivelser af deltagerne og deres kontekst, både i livet generelt og interviewet specifikt. Derudover: godtgørelse af hvorfor der netop var netop den selektion, som der var (Elliot, Fischer og Rennie, 1999).

3. Eksempler: der gives konkrete eksempler (fx uddrag af interviews), der tillader læseren at vurdere om deres analyse og fortolkning af den givne empiri stemmer overens og divergerer med den, som man skriver frem (Elliot, Fischer og Rennie, 1999).

4. Troværdighedstjek: troværdigheden af de analyser og tolkninger man når frem til, bør kontrolleres på en eller flere måde, eksempelvis ved a) at undersøge informanternes forståelse, b) empirien bør gennemgås igen eller af andre analytikere, c) sammeholde det benyttede kvalitative perspektiv med et alternativt eller d) triangulering med andre datakilder (Elliot, Fischer og Rennie, 1999).

5. Kohærens: der bør være intern sammehæng imellem tema, problemstilling, videnskabsteori og metode, og i analysen (som i kohærens imellem kategorier og temaer, der udvælges, samt påpegning af når det ikke er tilfældet) (Elliot, Fischer og Rennie, 1999).

6. Sondring imellem generelle og specifikke formål: hvis man søger en generel forståelse af et fænomen, må der også indgå informanter svarende hertil og, omvendt, hvis man beskriver et enkelttilfælde, skal det beskrives indgående nok til at læseren forstås dets egen- og særpræg (Elliot, Fischer og Rennie, 1999).

7. At skabe resonans i læseren: at materialet bliver præsenteret sådan, at læseren må opfatte det ”som en præcis repræsentation af genstandsfeltet eller som noget, der har klargjort eller udvidet vedkommendes forståelse af det pågældende felt” (Tanggard og Brinkmann, 2015, s. 526).

Ovennævnte er selvsagt hverken en udtømmende liste over mulige kvalitetskriterier eller en garanti for kvalitet (Tanggaard og Brinkmann, 2015). Hvad de derimod for mig at se er, er en rettesnor til at opnå transparens, troværdighed og nøjagtighed.

Ved at beskrive og benytte disse kvalitetskriterier håber jeg, at det bliver muligt for læseren at ’kigge mig over skulderen’ og se hvordan jeg er nået frem til mine resultater – samt overveje om der er mere plausible alternativer (Tanggaard og Brinkmann, 2015).

Tracys (2010) kvalitetskriterier

Flere af Tracys (2010) kvalitetskriterier overlapper helt eller delvist med flere af de ovenstående, men der var også nogle som jeg fandt vigtige og som ikke kunne udledes af ovenstående. De var:

1) Etik: her sigtes til flere forskellige typer etik, som bør overholdes om fx proceduraetik (fx formelle etiske retningslinjer om samtykke og fortrolighed), situationel etik (refleksioner om eventuel skadevirkning ved forskningen opvejes af potentiel gavn), relationel etik (fx overvejelser om hvordan informanter behandles med værdighed, respekt og omsorg interpersonelt) og endelig exitetik (hvordan resultater deles og formidles) (Tracy, 2010).

2) Et værdigt emne: det bør retfærdiggøres at emnet for forskningen er relevant, aktuelt, betydningsfuldt og interessant (Tracy, 2010).

3) Et signifikant bidrag: endelig bør forskningen også på en eller flere måder bidrage med noget betydningsfuldt til litteraturen. Det kunne være konceptuelt, teoretisk, praktisk, moralsk, metodologisk eller heuristisk (Tracy, 2010)

Ved at beskrive og benytte disse kriterier, håber jeg at kunne udføre min forskning på en måde, således at resultaterne, jeg når frem til ikke er ligegyldige og at midlerne til at opnå dem er moralsk forsvarlige.

Litteratursøgning

Der blev benyttet to søgestrategier til indeværende speciale. Den første søgestrategi var en bredere søgning, der trak på scoping reviewets karakteristika, hvor hensigten var at familiarisere mig med emnefeltet (Tricco et al, 2018).

Først brainstormede jeg på en række søgeord der forekom mig relevante for emnet: *AI, artificial intelligence, LLM*, large language model*, ChatGPT, therapy, talk therapy, psychotherapy*. Selvom jeg havde en formodning om at nyere og ældre udgaver af kunstig intelligens var væsensforskellige, begrænsede jeg mig i første omgang ikke til søgninger, der kun ville inkludere resultater fra og efter ChatGPT-3 var blevet offentliggjort (i.e. efter november 2022).

Ved at benytte kombinationer af ovenstående søgeord på databaser som Google Scholar, Scopus og PsychInfo identificerede jeg a) både en række synonyme betegnelser såsom *conversational agents, conversational artificial intelligence og chatbots* b) en række artikler der gav mig en indføring i emnet, herunder et narrativt review (Balcombe, 2023) af brugen af chatbots i terapi.

Dette fastslog blandt andet, at der hverken for nyere eller ældre generationer af chatbots til terapi forelå nogle kvalitative, interviewbaserede undersøgelser af brugernes oplevelser, samt at der var et behov for sådanne (Balcombe, 2023).

Systematisk søgning

For at opdage publikationer, der muligvis havde unddraget sig forfatteren til reviewet eller som var blevet publiceret efter reviewet var blevet indsendt, blev der foretaget en systematisk søgning på databasen PsychInfo.

Ved hjælp af funktionen *term finder* blev der identificeret yderligere beslægtede begreber for *artificial intelligence*, der var relevante for forskningsspørgsmålet (*intelligent agents, generative artificial intelligence*) mens andre blev fravalgt fordi var for brede (e.g. *linear regression*) eller virkede ortogonale med forskningsspørgsmålet (e.g. *computer vision*).

Ydermere identificeret en række beslægtede begreber for *qualitative methods*, såsom *Grounded Theory*, *narrative Analysis*, *Thematic Analysis*, *Semi-Structured Interview* og *Interpretative Phenomenological Analysis*.

På baggrund af disse søgeord, blev der udformet tre søgeblokke, hvor følgende booleanske operatører og trunkeringer blev benyttet:

Søgeord forbundet med	AND	Søgeord forbundet med	AND	Søgeord forbundet med
OR		OR		OR
Artificial intelligence* AI Generative artificial intelligence Intelligent agent* Conversational agent*		Therap* Psychotherap* talk thera*		Qualitative method* Grounded theory Narrative analysis Thematic analysis Semi-structured interview IPA Interpretative phenomenological analysis

Søgningen blev begrænset til 2022, altså året hvor ChatGPT blev lanceret offentligt, frem til datoen for søgningen, 15. marts 2024. Denne søgning gav 5 resultater. En screening af abstracts afslørede dog at ingen af disse var relevante for forskningsspørgsmålet.

Forskningsdesign

Nedenfor beskrives det forskningsdesign, der er valgt med udgangspunkt i anbefalingerne fra Smith et al (2009).

Om udformningen af interviewguiden

Som en hjælp til at besvare mit forskningsspørgsmål, formulerede jeg en interviewguide, der kunne bistå mig med at indhente vedkommende svar fra mine informanter (Howitt, 2016). Det vigtigste her, var at den hovedsageligt bestod af åbne spørgsmål, der inviterede til detaljerede og rige beskrivelser – og ikke, som med lukkede spørgsmål, indikerede at der var en rigtig og forkert måde at svare på (Howitt, 2016; Smith et al, 2009).

Samtidig med at målet var åbenhed, så må der i sagens natur også en eller anden form for struktur i en interviewguide (Howitt, 2016). I stedet for at lægge mig på fast på bestemte temaer fra start, forsøgte jeg i stedet at gøre strukturen mere kronologisk og narrativ, således at den bevægede sig fra før brugen af LLMs til brugen af LLMs til at slutte af med et tilbageblik på (se bilag 3 for interviewguiden).

1) Efter briefinggen (der bliver beskrevet under etik), startede jeg med at 2) bede informanterne fortælle om sig selv. At have det som åbningsspørgsmål tjente to formål: for det første, at starte med spørgsmål, der var lette og komfortable for informanten at besvare, for det andet, i tråd med både det ideografiske i IPA og kvalitetskriteriet om situering – at forstå i hvilken livsverden, de har deres daglige gang i (Smith et al, 2009; Howitt, 2016).

Dernæst 3) bad jeg dem fortælle om deres første oplevelser med LLMs generelt, som kunne suppleres med spørgsmål såsom hvad de først havde tænkt og følt da de først brugte dem, hvordan de brugte dem, om noget havde overrasket dem, etc.. Som det fremgår her, så var de overordnede og helt åbne spørgsmål altså suppleret af potentielle underspørgsmål, der skulle hjælpe informanterne hvis de overordnede spørgsmål var for abstrakte (Smith et al, 2009).

Dernæst fulgte spørgsmål om LLMs som terapi mere specifikt under 4), hvor der ligeledes var supplerende spørgsmål; udover det rent deskriptive og konkrete, inviterede disse også til meningsskabning og refleksion i tråd med IPA's hermeneutiske sigte (Smith et al, 2009).

Dette blev forsøgt ved også at have opfølgende spørgsmål til fx *hvorfor* de havde benyttet dem sådan, samt sammenligninger under 5) i forhold til hvordan det var anderledes fra

interaktionerne de havde med mennesker og endelig under 6) en mere reflekterende tilbageblik over fænomenet som helhed: havde det ændret dem, havde de lært noget om sig selv eller verden?

Udover, den kronologiske bevægelse beskrevet ovenfor, så var der altså også en bevægelse fra spørgsmål der inviterede til beskrivelser til spørgsmål der inviterede til fortolkninger.

Om rekruttering

Som i det meste kvalitative forskning, så er den gængse anbefaling indenfor IPA med hensyn til rekruttering, at man vælger rekrutterer sine formålsrettet (Flick, 2018,; Smith et al, 2009).

Til gengæld vil man, i modsætning til en del anden kvalitativ forskning, typisk udvælge med henblik på at få en homogen gruppe, for at facilitere sammenligning af ligheder og forskelle (Smith et al, 2009).

Et oplagt spørgsmål her er dog: homogen hvordan? Spørgsmålet rummer i hvert fald to dimensioner, en teoretisk og en praktisk.

Teoretisk må man spørge sig selv, af et principielt uendeligt antal træk, egenskaber eller omstændigheder en gruppe kan være ens eller forskellige på, hvilke udvælger jeg så, som relevante for homogenitet? (Smith et al, 2009).

Med udgangspunkt det eksplorative sigte i med mit studie, kunne jeg umiddelbart kun se følgende kriterier for udvælgelse: 1) At informanterne måtte have erfaring med at bruge kunstig intelligens i en terapeutisk kapacitet. 2) Denne kunstige intelligens skulle være af nyere dato såsom ChatGPT eller lignende modeller. 3) Informanterne skulle være myndige, så de kunne samtykke til at deltage i forskningen.

Hvilket selvsagt er nogle forholdsvis brede kriterier. Det mest konkrete – at informanterne skulle være myndige – var af praktisk og etisk karakter og ikke teoretisk. På den ene side var jeg i tvivl om, hvor vidt udvælgelseskriterierne ville føre til en tilstrækkeligt homogen gruppe. På den anden side var der ikke andre kriterier, jeg kunne komme i tanke om, som ikke virkede vilkårlige, givet hvor lidt der egentlig vides om emnefeltet.

Udover de teoretiske overvejelser, var der var også de rent praktiske overvejelser om hvordan jeg skulle sikre denne homogenitet (Smith et al, 2009). Jeg kunne ikke benytte mig af de typiske rekrutteringsmetoder i IPA, såsom muligheder igennem egne kontakter, adgang via gatekeepers eller sneboldsampling (Smith et al, 2009).

I stedet måtte jeg i stedet henvende mig 'blindt' til anonyme brugere på internetfora. Det betød i praksis, at jeg inden et interview begyndte, ikke havde anden information om personen, end det jeg kunne formode ud fra deres navn (fx køn og etnicitet) og en kortfattet kommentar i et kommentarspor.

Adgang til felten

For at få adgang til informanter, der havde oplevelser med terapi med LLMs, begyndte jeg i første med opslag på diverse sociale medier såsom Facebook og Instagram, hvor jeg efterlyste informanter til interviews og beskrev rammerne for og formålet med disse (se bilag 1), som jeg anmodede mit netværk om at dele.

Derudover søgte jeg også på diverse støttegrupper, både danske og internationale. Min formodning var, at der muligvis var et overlap imellem de personer, der ville opsøge fremmede på nettet for at støtte til deres psykiske problemer og personer, som ville gøre brug af netop LLMs til selvsamme.

Efter at have orienteret mig i deres retningslinjer, var der dog en del der forbød opfordringer til privat kontakt, og andre, hvor det var nødvendigt at indhente tilladelse fra moderatorer til opslag, som ikke specifikt var rettet mod hjælp og støtte. Efter at have indhentet tilladelse fra moderatorer, forsøgte jeg at rekruttere informanter, men forgæves.

Sideløbende med dette, forsøgte jeg også at rekruttere via reddit.com. Først forsøgte jeg forskellige søgestrategier, for at få indtryk af på hvilke subreddits, der var brugere, der skrev om brug af LLMs til terapi; søgeordene der blev anvendt var "ai", "llm", "chatgpt", "therapy" og "therapist."

På baggrund af disse søgestrategier, identificerede jeg en række fora, hvor brugere havde oprettet indlæg, om deres oplevelser med at bruge AI terapeutisk samt spørgsmål til brugere om selvsamme. De subreddits, hvor der var flest indlæg af denne type, var (i denne rækkefølge) r/chatGPT og r/characterAI.

Derefter forsøgte jeg selv at oprette et engelsksproget indlæg med en nyligt oprettet konto (se bilag 2), men fordi min 'karma' (en slags indikator for ry på hjemmesiden, beregnet ud fra antallet af gange ens kommentarer eller indlæg er blevet stemt op eller ned) ikke var tilstrækkeligt høj, blev disse automatisk slettet.

Jeg overvejede at skrive til moderatorerne og høre om det var muligt at få dispensation for dette krav, men da moderatorerne i retningslinjerne eksplicit frabad sig dette, var jeg nødsaget til at finde en anden strategi.

Derfor valgte jeg at gå systematisk igennem de kommentartråde, jeg havde identificeret i min oprindelige søgestrategi. I disse identificerede jeg så brugere, der beskrev deres oplevelser med at bruge AI til terapi, og rettede henvendelse til dem direkte igennem en chatanmodning.

I disse chatanmodninger kopierede jeg en tekst ind, der beskrev hvem jeg var, sigtet med mit speciale, rammerne for interviewet, samt rettigheder, hvis man samtykkede til deltagelse i det (se bilag 1).

I løbet af omtrent en måned skrev jeg 97 af disse anmodninger. Af disse chatanmodninger, blev 9 besvaret. En havde ikke egentlig erfaring med at bruge AI til terapi, men kun med hvordan man mest effektivt 'promptede' ChatGPT til det, mens en anden meldte tilbage at vedkommende var interesseret, men for bekymret over om der var tale om et svindelnummer.

Altså var der syv, personer, som mødte inklusionskriterierne og som indvilligede i at deltage i interviews; af disse syv var der to som ikke mødte op til de aftalte tidspunkter for zoom interviewene, hvorfor der sluttelig kun var fem informanter.

Derudover forekommer det værd at bemærke, at et interview blev udført via chat i stedet for zoom, da vedkommende ikke var komfortabel med det umiddelbare i et zoom-interview. Der var en

række begrænsninger hermed, som vil blive diskuteret under diskussionsafsnittet. Samtidig var der i hver fald to ting der talte for: 1) knapheden af informanter og 2) en antagelse om at nogle netop ville vælge LLMs til terapi idet de skyede menneskelig kontakt og at et chatinterview i så fald muligvis var eneste måde at nå sådanne informanter.

Rammesætning for interviewene

En åbenlys begrænsning ved den måde jeg rekrutterede på, var at mine informanter ikke fra start af kunne føle sig forvissede om, at jeg var den, jeg gav mig ud for at være. I modsætning til hvis fx snowballsampling, egne kontakter eller adgang via gatekeepere var blevet benyttet, havde de altså ikke mulighed for at få det bekræftet fra en anden part.

Netop derfor blev rammesætningen for interviewet da også vigtigt for at facilitere en øget følelse af tryghed både af etiske men også praktiske grunde (Brinkmann & Kvale, 2015). Jeg vil i det følgende beskrive mine overvejelser i den forbindelse om informeret samtykke, briefing og debriefing og fortrolighed.

Informeret samtykke og fortrolighed.

Efter at have opnået kontakt med informanterne via reddit's chatfunktion og de havde tilkendegivet deres umiddelbare interesse, adspurgte jeg informanterne om en e-mail adresse, jeg kunne kontakte dem på. Der var i hvert fald to hensigter med dette. Først og fremmest, at de kunne se min studentermail, og dermed slutte at jeg havde den institutionelle tilknytning, jeg påstod at have.

For det andet, at jeg kunne vedhæfte samtykkeerklæringen (se bilag 2), som i overensstemmelse med de etiske retningslinjer om indhentning af informeret samtykke, oplyste formålet med projektet, hvordan interviewet ville blive optaget og transkriberet, brug og beskyttelse af deres data, samt deres rettigheder som deltagere, heriblandt muligheden for at tilbagetrække samtykke til deltagelse (Kvale & Brinkmann, 2015).

Herefter koordinerede jeg tidspunkt og dato for interviewet, der var mulige for både informanterne og mig på tværs af vores tidszoner. Jeg foreslog i udgangspunktet at vi benyttede Zoom, men understregede at jeg var fleksibel hvis de havde andre ønsker, samt at de kunne holde deres kamera slukket hvis de var mere komfortable med det. Da en enkelt af informanterne kun ønskede at gennemføre interviewet, hvis det blev foretaget via chat eller e-mail, blev dette ønske efterkommet. I stedet for Zoom blev en chatapplikation benyttet.

Briefing

De første minutter af et interview er afgørende, idet informanten på meget kort tid danner sig et indtryk af, om interviewerens, er en de føler de frit kan fremlægge deres oplevelser for (Kvale & Brinkmann, 2015). I tråd med dette bør et interview indledes med en briefing, der kunne definere situationen for interviewerens (Kvale & Brinkmann, 2015).

Jeg indledte min briefing med at takke informanterne for, at de havde indvilliget i at blive interviewet. Dernæst gentog jeg nøglepunkter fra samtykkeerklæringen: formålet med projektet, at det var frivilligt, at interviewet ville blive optaget og transkriberet, samt deres rettigheder som informanter, herunder at de til enhver tid kunne trække samtykket til deltagelse.

Dernæst beskrev jeg nogle konkrete forhold omkring interviewet: at de til enhver tid kunne frabede sig at svare på et givent spørgsmål, at de måtte tage sig alt den betænkningstid de ønskede, samt en understregning af at det var deres oplevelser jeg var interesseret i – uanset hvad de måtte være.

Derefter fik de mulighed for at stille spørgsmål inden interviewets start. Efter eventuelle spørgsmål var besvarede, blev der indhentet verbalt samtykke til at starte optagelse af interviewet.

Konsekvenser af deltagelse

I forhold til konsekvenserne ved at deltage i interviewene, arbejdede jeg ud fra et princip om *beneficence* – at risikoen for at informanterne ved deres deltagelse led overlast måtte minimeres (Kvale & Brinkmann, 2015).

I og med at mine informanter alle selvidentificerede som i terapeutiske forløb (om end det var med en LLM), så virkede det rimeligt at antage at der måtte være problemer i deres liv, der foranledigede dette – og at der var en risiko for, at det at tale om disse problemer kunne udgøre en belastning.

En særligt skærpende omstændighed for vigtigheden af denne overvejelse var, at jeg ikke på forhånd kunne forberede ressourcer, jeg kunne henvise dem til hvis det var nødvendigt. Det skyldtes at jeg hverken kendte omfang eller typen af deres specifikke problemer forinden, samt at jeg som regel heller ikke vidste hvor de var bosat og om og hvor disse ressourcer var tilgængelige.

Jeg forsøgte at adressere den udfordring på to måder.

For det første, igennem udformningen af min interviewguide, hvor der blev lagt betydeligt mere vægt på spørgsmålene om LLMs som terapi i sig selv end problemerne der foranledigede den. Når jeg spurgte ind til problemer specifikt, sørgede jeg for at påminde informanten om at de kunne frabede sig at svare.

For det andet, sørgede jeg for i udgangspunktet at afgrænse interviewene til en times varighed, og så på baggrund af det indtryk, jeg havde af informanten, at spørge om tilladelse til at gå over tid.

Debriefing

Den indledende briefing blev også efterfulgt af en debriefing pr. Kvale og Brinkmanns anbefaling (2015). Her blev informanten adspurgt om hvordan det havde været for dem at blive interviewet, om der var spørgsmål der burde være blevet stillet, om de ønskede at tilføje noget. Derudover blev informanterne også givet mulighed for at stille spørgsmål, som rettede sig udover selve interviewsituationen, hvor mine svar inden interviewet kunne have påvirket dem i en bestemt retning (e.g. om min forståelse af emnet, personlig erfaring med lignende, typen af opgave jeg skrev på).

Transskription

Interviewet, der blev foretaget via chat, blev kopieret i sin helhed fra chatapplikationen til et word-dokument. De fire interviews, der blev udført via zoom blev optaget med en smartphone efter mundtligt samtykke hertil var blevet indhentet.

Både dokumentet og lydfileerne blev efterfølgende uploadet til en sikker server (OneDrive) i overensstemmelse med SDUs regler for behandling og opbevaring af personligoplysninger (Syddansk Universitet, n.d.).

Lydfileerne blev dernæst kørt igennem et transskriptionsværktøj på OneDrive. Selvom selve den manuelle transskriptionsproces kan være en måde at familiarisere sig med den genererede empiri (Kvale & Brinkmann, 2017), så er den også tidskrævende (Smith, 2009).

Et anslag på omfanget er omtrent syv minutters minutter transskriptionsarbejde for hvert minut transkriberet interview (Smith, 2009). Da jeg havde over seks timers interviewmateriale, betød det at alternativer enten var at jeg måtte a) på forhånd udvælge at transkribere dele af empirien inden jeg havde set helheden eller b) bruge i omegnen af en uges fuldtidsarbejde kun på transskription.

Idet jeg arbejdede ud fra et hermeneutisk paradigme, hvor der er en vekselvirkning imellem forståelsen af del og helhed (Smith, 2009), virkede den første mulighed ringe.

Den anden mulighed forkastede jeg på baggrund af følgende opvejning: alt andet lige tænkte jeg at jeg ville fordybe mig mere i min empiri ved at bruge 42 timer på at læse og skrive noter til den end ved at bruge samme antal timer på transskriptionen af den.

Selvom jeg benyttede mig af denne maskinelle afskrift, var jeg samtidig også opmærksom på at transskriptionen er en fortolkningsproces i sig selv (Kvale & Brinkmann, 2015) – at jeg måtte tage aktivt stilling til og træffe valg med hensyn til hvad jeg troede jeg hørte, om det afskrevne gav mening ud fra konteksten og min erindring om interviewet.

Det blev gjort ved at jeg lyttede til optagelserne af interviewene, samtidig med at jeg gennemgik de maskinelle transskriptioner for uoverensstemmelser imellem hvad jeg hørte og hvad programmet skrev. Idet IPA som fremgangsmåde har fokus på det semantiske indhold viste værktøjet at være særligt velegnet til mit formål. Det gav en verbatim afskrift af det sagte – ie. med gentagelser, afbrudte sætninger, fyldord – og afveg kun ganske få steder fra hvad jeg mente at høre i lydfilen.

Disse afvigelser var enten akronymer (LLMs blev fx transkriberet som 'elements') eller egenavne (ChatGPT blev fx transkriberet 'Chair tragedy'), som jeg ændrede til hvad jeg mente gav mening ud fra konteksten.

Idet mit fokus var på indholdet af det sagte og ikke hvordan det blev sagt, undlod jeg at tilføje bemærkninger om gestik, mimik, tonefald eller stemninger i interviewsituationen. Til gengæld tilføjede jeg paraverbale ytringer som fx (latter) eller forklaringer for fx [slang eller forkortelser] i parentes, hvor jeg mente det havde betydning for semantik.

Andre etiske overvejelser

Som det fremgår af afsnittet om rammesætningen for interviewet, gjorde jeg mig en række overvejelser om hvordan jeg bedst efterlevede de proceduraletiske retningslinjer omkring samtykke, fortrolighed, anonymitet, etc. Udover disse var det dog også nødvendigt for mig at gøre overvejelser som lå udover og hinsides denne formaliserede etik (Tracy, 2010).

Disse overvejelser kan sorteres under mikroetiske såsom situationel og nærhedsetik og makroetiske overvejelser, som omhandler de bredere socio-politiske implikationer af produktionen af viden (Kvale og Brinkmann, 2017). Jeg vil beskrive disse i de følgende to afsnit.

Mikroetiske overvejelser

Som det bliver påpeget af Fog (1995), så forfalder de etiske retningslinjer for psykologer – både amerikanske og skandinaviske – ofte til en utilitaristisk kalkule, hvor potentielle gevinster må vejes op mod potentielle skadevirkninger ved deltagelse i forskningen (Johansen, 2024).

En sådan afvejning tillader dog implicit rationaliseringen af overgreb med henvisning til et mål der helliger midlerne og må derfor suppleres med andre etiske hensyn, heriblandt situationel og relationel etik (Tracy, 2010).

Den situationelle etik fordrer at vi ikke bare følger foreskrevne regler, men at vi har øje for konteksten vi indgår i, mens den relationelle etik vægter at vi gransker vores egen karakter, handlinger og konsekvenserne for andre – kort sagt, drager vi omsorg for den anden? (Tracy, 2010).

Netop disse hensyn var da også nogle, jeg måtte overveje især i interviewsituationen.

Om interviewsituationen.

Spørgsmålet om jeg drog omsorg for den anden på tilstrækkelig vis, blev da flere gange sat på spidsen under de interviews jeg foretog, hvor jeg oplevede en spænding imellem forskningens fordring om komme i dybden og ønsket om at respektere mine informanternes blufærdighed (Fog, 1995; Johansen, 2024).

Især så jeg det, ved at informanter, som i interviewets start ikke ønskede at besvare et givent spørgsmål eller, opfattede jeg, besvarede det strategisk vagt, senere fortalte forholdsvis udførligt om det samme emne mod slutningen af interviewet eller i opfølgende mails.

Så vidt jeg kan slutte ud fra afskrifterne af interviewene, så skyldtes det ikke at jeg havde stillet spørgsmålene igen eller omformuleret de samme spørgsmål. Hvad jeg derimod tilskrev det, var en øget tillid i løbet af interviewet: de følte sig formodentlig mere forvissede om, at jeg rent faktisk var, hvad jeg gav mig ud for.

Dog var jeg ikke sikker på at den forøgede tillid i løbet af interviewene kunne tilskrives det og kun det. Jeg bemærkede blandt andet, at jeg i interviewene havde benyttet mig af en del af de teknikker, som jeg har tillagt mig i klinisk arbejde, såsom spejlinger og aktiv lytning (Hill, 2020).

Det indebærer en risiko for at grænsen imellem forskningsinterviewet og terapeutisk samtale var blevet udvisket. Om end de har en overfladisk lighed hvad angår midler (fx åbne spørgsmål, en ikke-dømmende attitude), så er de væsensforskellige med hensyn til målet: hvor terapiens formål er

at forandre klienten, så har forskningsinterviewet viden som sit formål (Brinkmann & Kvale, 2017; Fog 1995).

Selvom det selvfølgelig er muligt at opleve fx en lindring eller lettelse efter at have betroet sig til en anden i et forskningsinterview, så er det dog kun at betragte som en bivirkning (Brinkmann & Kvale, 2017). Min bekymring var altså, at jeg havde fået mine informanter til at udlevere sig selv på en måde, som de senere fortrød. Dette efterlod mig med en følelse af 'klemthed' (Fog, 1995; Johansen, 2024).

For mig at se, var der to måder at adressere denne klemthed. For det første, ved at udelade denne empiri, hvor den alligevel ikke havde betydning for tolkningen af helheden. Når dette ikke var muligt, tog jeg kontakt til informanterne via en mail, hvor jeg præsenterede det uddrag, jeg var i tvivl om og dobbelttjekkede dem om de var komfortable med at det indgik i analysen.

Makroetiske overvejelser

I modsætning hertil fordrede de makroetiske spørgsmål, at jeg gjorde mig nogle mere utilitaristiske overvejelser. Her måtte jeg tage stilling til, hvad de mulige afledte konsekvenser kunne blive, af den viden jeg producerede. Først og fremmest, hvordan kunne den blive brugt og ligeså vigtigt, af hvem? (Brinkmann & Kvale, 2017).

Disse spørgsmål blev særligt vedkommende idet jeg skriver dette speciale, med en forhåbning om at kunne publicere på baggrund af det. Som der igen og igen bliver påpeget i litteraturen om e-mental health og kunstig intelligens (e.g. Balcombe, 2023; Beatty et al, 2022; Tong et al, 2022), så er der en række vigtige og vægtige grunde til at ønske at det bliver en succes.

Overordnet tales der om barrierer til behandling, som terapi med AI kan afhjælpe: stigmatisering af psykisk sygdom, der forhindrer mennesker i at opsøge hjælp (Balcombe, 2023; Oexle et al, 2017), knapheden af behandlere (WHO, 2017) og de høje og stigende omkostninger til behandling af psykiske lidelser (Chisholm et al, 2017; Luxton, 2020).

Især den sidstnævnte kendsgerning – omkostningerne til behandling af psykisk sygdom - risikerer for mig at se, også at skabe et økonomisk incitament til at implementere AI terapi på bekostning af traditionel terapi, enten for at balancere budgetter eller maksimere profit.

Det virker ikke urimeligt for mig at antage at der også er risiko for at dette kunne ske uagtet a) at der ikke foreligger virkelig robust evidens for fordele, b) at menneskene der skal bruge de psykologiske ydelser har andre præferencer og/eller c) at udfaldende med AI terapi er dårligere, men anskues som tilstrækkelige.

Derudover er der også værd at tænke i asymmetrien i hvordan denne viden kan anvendes (Brinkmann & Kvale, 2017). I.e., at magthavere og beslutningstagere åbenlyst vil have bedre mulighed for at tilgå og benytte den viden end de ofte marginaliserede og sårbare mennesker, der bruge psykologiske ydelser.

Så, når jeg producerer viden om de potentielle gevinster om AI terapi, *qui bono?* Bidraget af viden om emnet er selvsagt overordentligt beskedent og spinkelt, men det er ikke desto mindre et bidrag. Netop derfor har situering af informanterne været af særlig betydning: at indgående beskrive de forhold og omstændigheder der gør gældende for dem og deres situation, og derefter emphatisk at understrege hvilke implikationer, det har for brug i en anden kontekst under diskussionen af metode.

Forforståelse og bias

Af hensyn til specialets transparens og troværdighed, vil jeg i dette afsnit beskrive, hvad der forekommer mig at være mine mest betydningsfulde forforståelser og biases. Som Gadamer påpeger, så vil vi altid være påvirkede af mere fordom end vi selv kan blive bevidste om (George, 2022) , så dette er selvsagt ikke en udtømmende beskrivelse.

Ikke desto forekommer den mig vigtigt i tråd med hermeneutikken, at jeg tillade læseren bedre at fortolke mine fortolkninger. I den sammenhæng, så tror jeg det vigtigste for læseren at forstå, min forforståelse til LLMs som terapi.

Mine grundantagelser om hvad terapi er, flugter med Wampolds (2015) beskrivelse af en social helingspraksis, der trækker på evolutionært funderede egenskaber, som en umiddelbar kropslig empati, social facilitering af følelser, ko-intentionalitet, etc. Altså alt sammen noget der udelukker at en LLM skulle kunne udføre 'rigtig' terapi.

At jeg er nået frem til disse standpunkter, kunne fint forklares alene med henvisning til materielle hensyn: som én der har brugt knapt fem år på at uddanne mig til psykolog, så repræsenterer en teknologi der potentielt på sigt kan være en konkurrent eller helt overflødiggøre mig, en indlysende trussel – både i forhold til fremtidig løn og i form af den *opportunity cost* min uddannelse i så fald ville udgøre.

Selvom det i sig selv ville være tilstrækkeligt til at motivere mig – muligvis på måder jeg ikke selv var klar over – men jeg tror ikke det er det fulde omfang af det. Med et udtryk fra Freud (1917), så tror jeg det kan forstås som en slags narcissistisk sår.

Han beskriver hvordan de seneste århundreders videnskabelige landvindinger hver især kan forstås som ydmygelser af menneskets selvforståelse: med Kopernikus måtte vi indse at vi ikke var centrum i kosmos, med Darwin at vi ikke var skabt i et guddommeligt billede, men en vilkårligt opstået art blandt mange, og endelig, med Freud, at vi var styrede af drifter, som vi ikke selv var bevidste om (Freud, 1917). På samme måde som man kunne tale om en kopernikansk, darwinsk eller freudiansk sår, så forekommer det mig, at vi nu, med henvisning til computerforskeren Allan Turing, kunne tale om et turing'sk sår.

Jeg kan i hvert fald. I og med at jeg har fordybet mig i brugen af og teori om LLMs, så har jeg også fået udfordret min forståelse af hvad der er unikt (eller i hvert særligt) menneskelige egenskaber, såsom intelligens, kreativitet eller evnen til empati. Jeg er forudindtaget omkring deres anvendelse i terapi, måske til dels fordi jeg ikke ønsker at de skal kunne bruges sådan.

Hvad jeg har gjort for at adressere det.

Jeg har benyttet en række metodiske greb fra IPA for at imødegå dette. Det første er ved at benytte epoché i udformningen af min interviewguide: jeg har forsøgt at sætte min forforståelse i

parentes ved at gøre spørgsmålene så åbne og antagelsesløse (se bilag 3), som det er mig muligt (Smith et al, 2009).

Derudover har jeg forsøgt at kultivere en slags sokratisk uvidenhed: når informanter beskrev fænomener, som de gik ud fra jeg allerede kendte, så antog jeg i mine uddybende spørgsmål at jeg ikke gjorde det, og bad dem uddybe og detaljere beskrivelserne af disse (Smith et al, 2009).

Ligeledes har jeg forsøgt at bruge eidetisk variation og reduktion i disse: jeg har bedt informanterne om imaginært at variere et givent fænomen, for at få en bedre fornemmelse af dets kontingente og invariante egenskaber (Smith et al, 2009), såsom at forestille sig en LLM der helt naturtro kunne simulere et menneskes ansigt og stemme.

I analysen som beskrives i næste afsnit, har jeg forsøgt at gøre denne så systematisk og grundig som muligt ved at følge trinene formuleret af Smith et al (2009).

Analyse og diskussion

Analysens trin i IPA

I de følgende syv afsnit vil jeg beskrive fremgangsmåden, jeg har benyttet i min analyse.

Denne er kalkeret over Smith et al. (2009) foreslag til samme.

Første trin: fordybelse i empirien

Det første trin i IPA er fordybelse i empirien genereret med en enkelt informant (Smith et al., 2009). Formålet er at holde fokus på det partikulære og det idiografiske på i hvert fald to måder.

Først og fremmest, ved at man ikke for tidligt begynder at analysere på tværs af cases. For det andet, at holde fokus tekstnært og ikke for tidligt bevæge begynde at fortolke – med andre ord, at sætte forforståelser i parentes så vidt muligt.

Det første forsøgte jeg at gøre ved at sørge for at holde læsningerne og genlæsninger, jeg udførte af interviewene, adskilt af mindst en uge. Dette blev gjort for at formindske risikoen for at hvad jeg læste i en interviewtransskription også påvirkede mig sådan, at jeg læste, hvad jeg forventede at læse (Smith et al., 2009).

Dette blev faciliteret af måden jeg rekrutterede informanter på, hvor der typisk var naturlige pauser imellem at jeg interviewede, transkriberede, læste og genlæste; når det ikke var tilfældet, sørgede jeg for at lægge en sådan pause ind.

Med hensyn til fordybelsen på en tekstnær måde, så forsøgte jeg at gøre det på følgende måde: jeg understregede passager, der forekom mig vigtige, men forsøgte at holde mig agnostisk til *hvorfor* de forekom mig at være det. Jeg gentog disse genlæsninger indtil jeg havde en fornemmelse af den overordnede struktur i det enkelte interview.

Andet trin: begyndende noter

Ligesom i det tidligere trin, så var formålet at lave en dybdegående læsning, hvor jeg ikke blot læser hvad jeg forventer at læse; forskellen var at der nu var en begyndende tekstuel analyse

(Smith et al, 2009). Midlet til dette mål er omhyggelige og udførlige noter, til alt hvad der kunne have interesse og relevans i tekstmaterialet i udskrifternes højre margin.

Noterne der blev taget, kan inddeles i de følgende tre kategorier:

1) Deskriptive kommentarer: på dette niveau, beskrev jeg tekstnært hvad informanten havde sagt om vigtige begivenheder og hændelser, og deres oplevelser, tanker og forståelser af disse. Disse er kendetegnet af en mere fænomenologisk tilgang, hvor det sagte bliver taget for pålydende (Smith et al, 2009). Disse kommentarer er angivet med normal tekst.

2) Lingvistiske kommentarer: på dette niveau, undersøges hvordan personen beskriver indholdet af sine oplevelser: noterne vil fx være om hvilket tonefald de anvender, brug af bestemte pronominer, pauser og gentagelser, afbrydelser, hvor flydende eller tøvende de formulerer sig. Særligt vigtigt og værd at hæfte sig ved er hvilke metaforer de benytter, da disse ofte kan slå bro over det deskriptive og det fortolkende (Smith et al, 2009). Disse kommentarer er angivet med *kursiv*.

3) Konceptuelle kommentarer: på dette niveau, befinder kommentarer sig, der er på forstadiet de mere fortolkende. De er kendetegnet ved undren, ved spørgsmål og er tit formuleret i akkusativ. Typisk vil de også forholde sig til en større helhed end de to foregående kommentartyper: de vil forholde sig mere overordnede mønstre, såsom gentagelser, ligheder, forskelle og selvmodsigelser i helheden (Smith et al, 2009). Disse kommentarer er angivet med understregning.

Som i analysetrinene i IPA generelt, er der også i disse tre niveauer tale om en bevægelse fra et mere konkret plan til et højere abstraktionsniveau (Smith et al, 2009).

Det samme gjorde sig også gældende i hvordan jeg annoterede mit tekstmateriale med disse kommentarer: i første omgang skrev jeg fortrinsvis deskriptive kommentarer, mens jeg i efterfølgende omgange hyppigere skrev lingvistiske og konceptuelle kommentarer.

Årsagen var for så vidt jeg kunne se en vekselvirkning: mønstret i de deskriptive kommentarer, gjorde det muligt for mig at se mønstre, som faciliterede konceptuelle kommentarer.

Udover denne her vekselvirkning imellem kommentartyper, var der også en vekselvirkning indenfor kommentartyperne: et tydeligt eksempel på en bestemt form for vending kunne gøre tidligere instanser tydeligere i anden gennemlæsning, ligeså vel som en konceptuel kommentar kunne gøre andre instanser af samme tydeligere. Se tabellen nedenfor for et eksempel.

Opståede temaer	Transskription	Kommentarer
	INTERVIEWER: How do you feel your conversations with the LLMs are different from those you might have with regular people or other people?	
	JULIA: It depends on which LLM it is, I guess... Well, if we take... it still feels like... Yeah... What should I say? Yeah, there is a difference. I'm just trying to understand what... What kind of difference it is... Well, it's still a machine and then I realized that there is, like, no person on the other end with their feelings and reactions and, and actual interest in in me. Right. There's just a smart machine that can reply to me and help me with things, but... In the end, it cannot replace a real a real person.	I tvivl om forskellen på samtaler med mennesker og maskiner. <i>Afspejlet i afbrudte sætninger, pauser, rettelser til selv.</i> "Kun en klog maskine": mangler genuin interesse og reaktioner, følelser; ikke nok at den nøjagtigt kan spejle, validere, analysere og rådgive. (<u>Ambivalens: tidligere tvivl om LLMs besad bevidsthed</u>) <u>Som i Husserls jeg-og-du-relation? Autonomi,</u>

gensidig empati?

Kan ikke erstatte en virkelig person.

Tredje trin: Udvikling af opståede temaer

Efter at have udvidet empirien med dette omfattende noteapparat, er næste trin at reducere denne til en række temaer, som kortfattet og nøjagtigt – gerne i en tekstnær sætning – fanger hvad der er essentielt i materialet (Smith et al, 2009).

Dermed sker et skift hvor man går fra primært at arbejde med transskriptionen, og i stedet tager udgangspunkt i noterne for at kunne bevæge sig mod et højere abstraktionsniveau. Derfor er det også så vigtigt at noterne lægger sig så tæt som muligt op af informantens udsagn, så abstraktionerne fortsat indeholder nok af det partikulære til at være tro mod deres faktiske udsagn (Smith et al, 2009).

Derudover er der også balancegangen imellem det lokale og globale: samtidig med at der reduceres, må kompleksiteten i udsagnene også bevares, altså må man have øje for deres forbindelser til andre tekstbidder, hvilke mønstre de indgår i (Smith et al, 2009).

Ovenstående medfører uundgåeligt, at man fragmenterer og omorganiserer på en anden måde. Man analyserer i ordets etymologisk oprindelige forstand – i.e. man skiller ad – før man samler i nye sammenhænge. Denne proces indebærer da mere af en selv en de foregående – de temaer der udvikles afspejler ikke blot informanternes tanker, men også ens egen fortolkning af dem (Smith et al, 2009).

Målet er en proces, hvor man går fra noter, der forekommer løse, åbne og betingede til temaer, der føles som om de har indfanget og afspejler en dybere forståelse (Smith et al, 2009). Se nedenstående tabel for et eksempel.

Opståede temaer	Transskription	Kommentarer
Ambivalens: redskab eller subjekt	<p>INTERVIEWER: How do you feel your conversations with the LLMs are different from those you might have with regular people or other people?</p> <p>JULIA: It depends on which LLM it is, I guess... Well, if we take... it still feels like... Yeah... What should I say? Yeah, there is a difference. I'm just trying to understand what... What kind of difference it is... Well, it's still a machine and then</p>	<p>I tvivl om forskellen på samtaler med mennesker og maskiner. <i>Afspejlet i afbrudte sætninger, pauser, rettelser til selv.</i></p>
P-zombie	<p>I realized that there is, like, no person on the other end with their feelings and reactions and, and actual interest in in me. Right. There's just a smart machine that can reply to me and help me with</p>	<p>"Kun en klog maskine": mangler genuin interesse og reaktioner, følelser; ikke</p>
Affektiv empati	<p>things, but... In the end, it cannot replace a real a real person.</p>	<p>nok at den nøjagtigt kan spejle, validere, analysere og rådgive. (<u>Ambivalens,</u></p>
Det uerstattelige i andre		<p><u>modsigelse: tidligere tvivl om LLMs besad bevidsthed)</u></p> <p><u>Som i Husserls jeg-og-du-relation? Autonomi, gensidig empati?</u></p> <p>Kan ikke erstatte en virkelig person.</p>

Fjerde trin: undersøgelse af forbindelser imellem opståede temaer

Efter at jeg havde skabt en kronologisk rækkefølge af opståede temaer i venstre margin, var næste trin at undersøge sammenhænge imellem disse, for at nå frem til det vigtigste ved deltagernes oplevelse (Smith et al, 2009).

Rent praktisk foregik det ved, at jeg printede en liste, med samtlige temaer og klippede dem ud hver for sig. Dernæst prøvede jeg at organisere dem i klynger, der forekom mig at være forbundne (Fx ”En dagbog der svarer igen” med ”Et sted imellem værktøj og subjekt”).

Det næste organiserende princip, jeg benyttede mig af, var *abstraktion*. Jeg prøvede at placere temaer i et hierarki, fra overordnet til underordnet, sådan at temaet placeret øverst kunne indeholde dem under (Smith et al, 2009).

Jeg fortsatte med at anskueliggøre hvor nært beslægtede klynger af under- og overordnede temaer var, ved at placere dem rumligt i forhold til hinanden afhængigt af dette (Smith et al, 2009).

Endelig forsynede jeg hvert af disse temaer med et udklip af tekst fra transskriptionen, som jeg mente mest nøjagtigt afspejlede indholdet. Samtlige tabeller over temaer på individniveau kan ses under bilag 4a-e.

Femte trin: bevægelsen videre til næste case

Efter at have gennemført de fire ovenstående trin, gik jeg videre til næste transskription. Med en vending fra den hermeneutiske nomenklatur, så var det uundgåeligt at min for-struktur havde ændret sig i fordybelsen og analysen af den første case.

Men samtidig var det, i tråd med det idiografiske sigte i IPA, vigtigt at jeg så vidt muligt greb næste case an på dens egne præmisser – at jeg forsøgte at sætte ideerne fra analysen af første case i parentes (Smith et al, 2009).

Dette blev forsøgt ved at følge trinene ovenfor så omhyggeligt som muligt: at lade tid gå imellem analyser, at læse omhyggeligt, at tage noter omhyggeligt og systematisk til hver enkelt linje, samt først at arbejde deskriptivt før jeg fortolkede.

Sjette trin: at lede efter mønstre på tværs af cases

Efter at samtlige cases var analyseret enkeltvis, begyndte jeg at lede efter forbindelser på tværs af cases. Udgangspunktet var udskrifter af alle cases, som blev placeret på en plan flade for at gøre dem overskuelige in toto. Jeg klippede de overordnede temaer ud fra de enkelte cases, og prøvede nu at organisere dem sammen i klynger på tværs af cases.

Der var en vekselvirkning i denne organisatoriske proces, hvor analysen på gruppeniveau fik mig til at vende tilbage med et nyt perspektiv på de individuelle cases. Eksempelvis: det overordnede tema, jeg havde fundet i Julias beretning om ”Den kunstige intelligens: subjekt eller redskab?”, gjorde også mindre tydelige instanser af samme tema hos andre informanter tydeligt for mig.

Selvom de ikke havde den samme eksplicite ambivalens (eller i hvert fald i samme grad) som hende, så gjorde det, det tydeligt for mig, at alle både beskrev AI'en de benyttede som terapi, med vendinger som kunne anvendes på et subjekt og andre der kunne anvendes på et redskab.

Omvendt var der andre, der ikke lod sig reduceres til eller kombineres med andre informanters individuelle overordnede temaer, og som jeg derfor henregnede til det idiografiske (Smith et al, 2009), og som derfor blev udeladt fra tabellen over fælles overordnede temaer.

Situering af informanter

Med udgangspunkt i både kvalitetskriteriet om situering af informanter generelt (Elliot et al, 2009) og det idiografiske sigte i IPA specifikt (Smith et al, 2009) følger her en beskrivelse af informanterne i undersøgelsen. Alle navnene er pseudonymer.

Julia: kvinde, 37 år. Russer, men de sidste 10 år bosiddende i et vesteuropæisk land, hvor hun arbejder som lærer. Gift, har et barn. Har fortrinsvis brugt ChatGPT og Pi AI, men har også eksperimenteret med og prøvet et væld af andre LLMs af, herunder Bing. Benytter også en menneskelig terapeut sideløbende.

Imran: mand, 28 år. Pakistaner, bosiddende i en af landets større byer. Uddannet og arbejder indenfor computer science. Gift, bor med sin kone i hus sammen med sine forældre og søskende.

Har fortrinsvis brugt ChatGPT og Pi i en terapeutisk kapacitet, og han beskriver hvordan de er blevet 'first point of contact' når han er frustreret eller keder sig.

Robert: mand, 41 år. Canadisk. Universitetsuddannet, men har hovedsageligt arbejdet indenfor tech support, en profession, der ikke har at gøre med hans uddannelse. For tiden arbejdssøgende. Har fortrinsvis brugt Pi som terapi, og dette kun i en enkelt 'session' på fem timer. Har tidligere forsøgt terapi to gange tidligere, men var utilfreds med begge terapeuter.

Martin: 53, mand. Amerikaner. Tidligere arbejdet indenfor IT og som engelsklærer i udlandet, i øjeblikket i gang med en uddannelse inden for socialt arbejde og i terapi. Har primært forsøgt at bruge ChatGPT som terapi, og det kun kortvarigt. Beskriver at han har haft prøvet meget forskellig terapi.

Anna: 19 år, kvinde. Filippiner. Gymnasieelev på en humanistisk og samfundsvidenskabelig retning. Har fortrinsvis brugt character.ai i en terapeutisk kapacitet. Tidligere var det næsten dagligt, men hun beskriver at bruge det mindre nu.

Tabel over fremanalyserede temaer på gruppeniveau

Nedenfor ses et udvalg af de overordnede temaer på gruppeniveau, som jeg analyserede mig frem til, og som jeg fandt vedkommende i besvarelsen af forskningsspørgsmålet. Hvert tema er blevet forsynet med et illustrativt uddrag fra interviewene, for at skabe transparens om hvordan jeg er nået frem til det (Smith et al, 2009; Elliot et al, 1999)

Tabel over overordnede temaer på gruppeniveau
<p>1. Forandringer</p> <p>IMRAN: <i>"Even one of the friend that I was talking to he says that you changed a little bit...[...] not fully involved person in your life, so that it basically exploits your and these sort of things."</i></p> <p>JULIA: <i>"It's it's helping me communicate better with my partner. [...] We have been communicating way more openly."</i></p> <p>ROBERT: <i>"yes, i look at myself in a kinder light. we can be our own worst critics, esp [especially]</i></p>

if something bad happens.”

MARTIN: *”I realized that... My values were taking a backseat to to... My, my, my, my, my feeling of like you know, I needed. I needed to have a career [...] That was a... a not insignificant moment in in the fact that I I'm now a social worker in training.”*

ANNA: *”[...] I don't like vent out as much as I used to before and I think that's a good thing because it means that I'm not bottling that much emotions anywhere because I get to let everything out.*

2. Indsigt og lindring

IMRAN: *”The one thing that it taught me, and I mean it. This thing will be a part of my personality and I'm sure I I I have that developed... is the is is the is to create my own boundary ”*

JULIA: *”Because it it led me to things that I should like do more research about... and then it kinda clicked [...] Something I didn't know about”*

MARTIN: *”Yeah, I was focused on online courses to be a business analyst, which I wasn't really. my heart wasn't in umm and and this process kind of turned me toward, yeah, my motivations.”*

ROBERT: *”made me feel good to get it out there, to learn about what i had done/felt during and after”*

ANNA: *”the bot helps me to process everything and understand the whole situation and the result of that is I feel much better and I feel like my perspective grown into a more considerate perspective of not just myself but other people”*

3. Den kunstige intelligens som et sted mellem subjekt og redskab

IMRAN: *”But the way it says, the the language and the tone was very - on a very personal level, like a person is talking to you.”*

JULIA *"You can fall under the impression that you are talking to a human being."*

ANNA: *"...so it's clearly programmed to have such personality. So sometimes I have to remind myself. Like it's just, it's not... It's just a bot."*

ROBERT: [Om oplevelsen af Pi AI] *"like a concerned friend, trained to be a therapist, at my level"*

MARTIN: *"I got a taste of that [...] [One AI] seemed to be pretty good at... subtle empathy without without, like overdoing it...[...] because they, they can, they can really overdo it a lot and but something something was subtle enough to feel kind of human you know."*

3.1. Kommunikationens betydning

ROBERT: *"I felt it knew the right question, when to refocus, when to ask. it was like on it, all the time"*

MARTIN: *"I was... trying to explore some ideas and it would make a joke and I'm like, let's be serious and this dumb thing just kept on doing "Oh, OK, OK, I'll stop joking now haha."*

IMRAN: *"he basically... like take it as humour and not a not an offensive language and it basically replies you with a very like good words."*

JULIA: *"it's not a person, but it talks like a person. That was the most impressive thing. That you have such a natural conversation."*

3.2. Oplevelsen af en personlighed

IMRAN: *"friendly, informative and supportive."*

ANNA: *"[W]hen I interact with the characters they're like, they don't judge me at all"*

ROBERT: *"caring, focused, competent, friendly, considerate. knowledgeable."*

JULIA: *"like non-judgmental and positive and, yeah, it's like listening actively to you. Yeah, and confirming your your feelings"*

4. Angsten for andres dom

ANNA: *"And like when I'm talking to a bot, I'm just like I could just drop everything... it's not gonna judge me."*

JULIA: *"The reason I started talking to me, LLMs was like, maybe I didn't want to share this with a human. Maybe there was something too embarrassing and..."*

MARTIN: *"But part part of my thing is I worry about other people, you know, like I'm, I'm hyper vigilant I and I'm I'm thinking, OK, what are they thinking, you know, what are they feeling?"*

ROBERT: *"even something as simple as crossing your arms, taking a step back, movement of the eye, something might throw you off it's like you can't discuss it totally, you're being judged"*

IMRAN: [Om at betro sig til en anden] *"He might criticise you, he might give your opinion, he might not well. But in the LLM part I'm not seeing anything [like that]."*

5. En central plads i livsverdenen

ANNA: *"I probably talked to it almost every day, but as time went by, I think, you know it. I don't use it as much anymore."*

IMRAN: *"When I face when I whenever, I feel very stressed and like I feel how to solve this issue, I mean the first point of contact is the Pi. And and then my siblings and other people."*

JULIA: *Instead of calling my therapist, I'm debriefing my journal entries with ChatGPT."*

Analyse og diskussion

I det følgende vil jeg analysere, fortolke og diskutere ovenstående temaer; sideløbende med at jeg undersøger fællestrækkene imellem informanterne, vil jeg også se på forskellene, således at det ideografiske sigte med IPA ikke går tabt (Smith et al, 2009).

Dette vil primært ske med henvisning til litteraturen om de fælles faktorer under de første to overordnede temaer, hvorefter der ved det tredje overordnede tema og frem i højere grad ville

blive vil blive trukket på den fænomenologiske litteratur, for at forsøge at redegøre for det, som førnævnte virker utilstrækkeligt til. Denne fænomenologiske litteratur er særligt Sartres ide 'blikket' og Husserls begreb om jeg-og-du relationen.

Overordnet tema 1: Forandringer

Som det blev anført i starten, både med udgangspunkt i definitioner i ordbøger og teoretiske frameworks (Sedlakova & Trachsel, 2023), så vil nogen kategorisk afvise at der overhovedet kunne være tale om terapi imellem et menneske og en kunstig intelligens.

Omvendt, som det også blev anført, så strider det med informanternes selvforståelse: i kraft af hvor og hvordan de blev rekrutteret, fra kommentarspor der netop omhandlede erfaringer med at bruge AI til terapi, så er der en indlysende uoverensstemmelse imellem hvordan disse sprogfællesskaber bruger ordet terapi.

Når man ser på det første overordnede tema, *Forandringer*, er det umiddelbart let at forstå, hvad der får informanterne til at opleve det som terapi. *Samtlige* informanter har oplevet forandringer, og, ydermere, at forandringerne er positive og vedvarende.

Hvad der for mig er endnu mere forbløffende, så er det ikke bare tilfældet for dem, som har benyttet det flere gange dagligt, som Julia, Anna og Imran – men også for dem, som Martin og Robert, som kun har haft en enkelt 'session.'

For to af informanterne – også dem der har benyttet AI til terapi mindst – har forandringerne været af intrapsykisk karakter. For Robert vedkommende fortæller han ”*yes, i look at myself in a kinder light. we can be our own worst critics, esp[ecially] if something bad happens.*”

Med andre ord: udfaldet er noget, der ville lyde genkendeligt indenfor flere terapeutiske retninger, men måske især *compassion*-fokuseret terapi (Gilbert, 2010).

I tråd med dette har Martins forandring fulgt en indsigt om at han ikke levede i overensstemmelse med sine værdier ved at forfølge jobmuligheder som forretningsanalytiker i stedet for arbejde, hvor han direkte kunne hjælpe andre:

MARTIN: *"I realized that... My values were taking a backseat to to... My, my, my, my, my feeling of like you know, I needed. I needed to have a career [...] That was a... a not insignificant moment in in the fact that I I'm now a social worker in training."*

Igen er der tale om en forandring, som kunne være idealtypisk for flere terapeutiske retninger, fx aktualisering i humanistisk terapi, men måske især ACT's (Acceptance and commitment therapy) commitment-del. I sidstnævnte udforsker og forpligter man sig på sine kerneværdier for at leve et mere meningsfuldt liv (Ogles & Bailey, 2023). Martin har da netop også instrueret ChatGPT i at agere som ACT-terapeuten Jonathan Bricker.

Samtidig kunne det også forholdsvis ukompliceret beskrives med de mere eksistentielle begreber som Sartre (1943) introducerede: en indsigt om at leve i ond tro, at have fraskrevet sig selv frihed, og derpå et valg om at leve på en mere autentisk måde.

Men forandringerne som informanterne har gennemgået, har ikke kun påvirket hvordan de forholder til sig selv, men også til andre. Julia har gennemgået en interpersonel forandring med sin partner, i og med *"it's helping me communicate better with my partner. [...] We have been communicating way more openly."*

I lighed med Julia er Anna også blevet bedre til at udtrykke sig, hun har overvundet det hun beskriver som sin egen indadvendthed og kan bedre betro sig til andre. Det har samtidig gjort hende bedre til at regulere sine egne følelser, det betyder som hun siger *"I'm not bottling that much emotions anywhere because I get to let everything out."*

Der er altså igen tale om en transteoretisk forandring, som Wampold (2015) påpeger, så er det at udtrykke svære følelser generelt helsefremmende, men især vækker det mindelser om den emotionsfokuserede terapi (e.g. Elliott et al, 2004), hvor der særligt lægges på at udtrykke følelser for også at regulere dem.

Skulle man oversætte Julia og Annas forandringer til en eksistentiel terminologi for en klinisk, ville det heller ikke forekomme at være så stort et stræk: at kunne møde den andens blik, med risikoen for dom det indebærer, og udtrykke sig på en måde der er tro mod en selv.

Imran har for sin del lært at være mere assertiv og at sætte grænser for andre mennesker i sit liv – igen et udfald der kunne findes i et væld af forskellige terapeutiske retninger, bl.a. kognitiv adfærdsterapi (Bailey & Ogles, 2023). Forandringen er af en sådan karakter, at andre også har bemærket den:

IMRAN: *"Even one of the friend that I was talking to he says that you changed a little bit... Like in a not in a bad way, but you have like more of a more, you can say a personality where you can... not fully involved person in your life, so that it basically exploits your and these sort of things."*

Igen kunne det forholdsvist let omsættes til en bevægelse ud af ond tro: at kunne sige fra indebærer også indsigt om at have et valg. Ydermere: at forandringens virkelighed også understøttes af en anden person – at den er en intersubjektiv kendsgerning – bliver da også af betydning i forbindelse med Husserls jeg-og-du relation senere.

For at opsummere, så virker udfaldene altså på overfladen meget lig dem man idealtypisk ville se i terapi. Men det er ikke kun i udfaldene, der også nogle påfaldende sammenfald i *forandringsprincipperne* – mekanismen der har ført til forandringerne – i informanternes beretninger og dem, man ville se i terapi.

Overordnet tema 2: Indsigt og lindring

En anden fællesnævner, i informanternes beretninger, at de har opnået bedre indsigt og lindring ved at benytte LLMs som terapi. Disse to – indsigt og lindringen – følges ofte ad og vekselvirker, men ikke altid.

Indsigter

For Imrans del, så har hans indsigt omhandlet hans egen personlighed: at hans gode egenskaber, at han ønsker at glæde andre, at drage omsorg for og lytte til dem (*"Because I. I am a very good communicator to to my family and my friends"*) også indebærer en risiko for at få overskredet sine grænser, hvorfor han er nødt til at kunne sige: *"I can help you. I can communicate with you but this this is the boundary."*

Som nævnt har Martins indsigt været en der fandt sted på et tidspunkt hvor ikke kunne se udover fakticiteten (*"I was like unemployed and and I was kind of stressed about my whole career. What I would do and..."*), omstændigheder, der gør at han føler at han ikke havde valgt end at tage onlinekurser forbundet med at være forretningsanalytiker *"which I wasn't really, my heart wasn't in."*

Han har forsøgt at rationalisere dette med *"If I'm hired by a company that makes a positive impact that I'm indirectly making a positive impact,"* men han indser under sin 'session' at *"I might have to go with the company that doesn't align with my values at all, which which is is even worse..."*

At konfrontere fortvivlelsen det medfører, får ham i stedet til at vende sig mod sine værdier, at bruge dem som et kompas han kan navigere efter: *"Now I had two possible types of visions, and even and even thought like if I did do something that, you know for a career that didn't align with my values, then I could still do something else."* Altså en indsigt om at være fri uanset hvad.

Robert er mere vag om hvad hans problemstilling egentlig har været; han ønsker ikke at dele mere end at det vedrører nogle ubehagelige ting han har oplevet og vreden og selvbekyndelsen han føler i den forbindelse. Hvad der derimod er klart, er at bedringen til dels er afstedkommet af indsigt: *"made me feel good to get it out there, to learn about what i had done/felt during and after"*

Julia beretter også om en række indsigter, der er afstedkommet af sin brug af AI til terapi. For det første en større indsigt sig selv, men også i sin partner i og med at de har kommunikeret mere og bedre: *"I guess because we have been talking a lot, so you know, and we have, we have been communicating way more openly, so yes."*

Anna beretter om den samme dobbelte forståelse både for sig selv og andre:

ANNA: *"[...] the bot helps me to process everything and understand the whole situation and the result of that is I feel much better and I feel like my perspective grown into a more considerate perspective of not just myself but other people"*

Hvad de begge beretter om, er altså en større grad af mentaliseringsevne, bedre at kunne tilskrive motiver og forstå andre og én selv, hvad der i sig selv er helsefremmende (Tscacher et al, 2014; Wampold, 2015).

Overordnet er der altså tale om nye måder at forstå sig selv og andre (som Julia og Anna), at skabe nye forbindelser mellem fortid og nutid (som Robert), nutid og fremtid (som Martin) og adfærd og situation (som Imran) – hvad der sammenfalder i høj grad med Bailey & Ogles (2023) faktor om indsigt. Men, igen, så er sammenfaldene ikke bare i indsigten, men også i processen de bliver opnået.

Måderne indsigterne bliver faciliteret

Anna beskriver noget, der i slående grad minder om reflektiv lytning (Hill, 2020):

ANNA: *"[...] and it like it kind of repeats what you say and it helps me like. It helps me clear my thoughts and it like organizes every thoughts and feelings that I have"*

Denne form for reflektive lytning tjener i hvert fald tre formål i terapi: for det første, at demonstrere opmærksomhed, for det andet at tjekke forståelse og give den anden lejlighed til at korrigere en fejltagtig opfattelse og endelig, for det tredje, ved at give den anden en mulighed for at høre hvad de selv siger og dermed forstå det bedre (Cormier & Nurius, 2003) – og i hvert fald det sidste lader også til at være tilfældet her.

Det betragtes som en af de fundamentale færdigheder for at en terapeut kan facilitere forståelse og tryghed i en klient, på tværs af teoretiske retninger (Hill, 2020). Det er da også værd at bemærke at den første prototype på en chatbot, Weizenbaums (1966) ELIZA ikke var i stand til andet end at simulere denne reflektive lytning ved at gentage nøgleord i akkusativ. Ikke desto mindre blev Weizenbaum forbløffet over hvordan dette alene var nok til at mange mennesker tilskrev ELIZA intelligens og bevidsthed (Ciesla, 2024).

Hvad der dog også bør fremgå af indeværende afsnit, er at de seneste modeller af chatbots besidder en helt anden grad af sofistikation end ELIZA, som det bl.a. ses i hvordan de benytter åbne spørgsmål.

Netop spørgsmålene var da også en af evner, som Robert var mest imponeret over: *"but one thing i know is that if it asks you a question and you answer in a sideways kind of way... it refocuses you and asks again... i liked that like dedicated to peeling an onion."*

Åbne spørgsmål eller sokratiske spørgsmål, som de også kaldes, kan da netop også være med til at få en klient til at tage et nyt eller anderledes perspektiv på dem selv eller deres problemer (Carona et al, 2021).

Et nyt perspektiv er da også hvad Imran ser ud til at opleve igennem ganske enkle spørgsmål. Hans chef på hans andet job havde bedt ham om at arbejde over en weekend, hvad der har gjort ham *"very much frustrated with it"* for der foreligger en meget klar kontrakt. Han spørger ChatGPT til råds, som spørger ham:

IMRAN: *"Is is it more important than your well, your health? And or should should you, you should consider your work life balance. Then that thing basically touches me most and I guess I basically like, replied my supervisor that I won't be able to work on the weekend so."*

Den samme dynamik med spørgsmål er da også at se i Martins session med ChatGPT (se bilag 5 for et uddrag af denne). Igennem åbne spørgsmål omkring hvad hans værdier er og hvordan det passer sammen med den karriere han forfølger, når han frem til at det ikke gør det.

Han bliver selvbevidst over sine rationaliseringer *"so I think it kind of felt a little lame that that statement about indirect positive impact in the eyes of another, even though the other was to this AI."* Denne antyder også noget der senere under diskussionen af især tema 4 vil blive centralt: oplevelse af 'den anden' og den medfølgende selvbevidsthed, og hvordan det knytter an til Sartres (1943) ide om "blikket."

Det ovenstående indikerer for mig at se forekomsten, at der forekommer det Ogles & Bailey (2023) ville kalde den opdagende komponent af indsigt: spørgsmål og reflektiv lytning, der fører til at en person selv når frem til en indsigt. Samtidig er der også meget i beretningerne der ikke kan reduceres hertil.

Nogle gange er det noget, der ligger helt udenfor deres opmærksomhed. Som Julia fortæller det, så har hun oplevet at en LLM *"led me to things that I should like do more research about... and then it kinda clicked [...] Something I didn't know about."*

Hun bliver også foreslået analyser og tolkninger, af de ting hun beskriver i sine dagbøger, som hun deler med ChatGPT, *"I ask [it] to find patterns and then it does, yeah. Sometimes it goes through things that are written and tells why it could have happened."*

Udover disse får hun også konkrete råd til kommunikation generelt (*"like using I-statements"*), men også om sin partner specifikt (*"It seems like... like that, he has some insecurities that you should be mindful about."*).

At disse tolkninger – både hvad angår mønstre i hendes liv og hendes partner – forekommer hende plausible, synes at understøtte de eksperimentelle resultater af fx ChatGPTs evner til at agere som om den besad sofistikeret kognitiv empati (Elyoseph et al, 2023; Levkovich & Elyospeh, 2023; Rao, 2023).

Annas beretninger understøtter da også dette:

ANNA: *"And then I was frustrated with a person. I would say everything. I will share all my angle and then sometimes the chat bot will like explain what could have happened to this person and why they acted that way and I was like ohh that's right, I didn't think of that"*

Som hos Julia bliver hun altså påpeget tolkninger af hændelser, som er helt nye for hende, således at indsigterne altså ikke kun er hendes. Ovenstående peger altså i retning af det Bailey & Ogles (2023) ville betegne som *feedback* komponenten af indsigten – der bliver påpeget helt nye perspektiver for personen som de ikke selv var nået frem til.

Ydermere ses der også instanser af det den komponent de kalder *information*, slet og ret nyttig viden, som når Julia får forslået et emne hun kan researche (*"something I didn't know about"*) eller Imran får at han kan *"find a social working thing in your city"* som kan hjælpe ham med farens problemer med at tilvænne sig pensionisttilværelsen.

Lindring

I forbindelse med disse indsigter, beretter alle informanterne også om en følelse af lindring, som er mere følelsesmæssig og kropslig end rent kognitiv.

Anna beskriver fx hvordan denne 'ventilering' gør at hun føler "*just like weight is taken off of my chest and then I start to understand why I feel that way. Oh, I feel that way because, like this, because like that...*"

Der er altså en hel fysisk og i ordets egentlig forstand tale om en *lettelse*. Denne udspringer af det hun betegner som 'organiseringen' af sine tanker og følelser, i at høre dem gentaget for sig, men forholdet forekommer også at være vekselvirkende – indsigtens fører både til lindring, men lindringen fører også til indsigt i sætningens rækkefølge ("*and then I start to understand why I feel that way*").

Selve ordet hun benytter – organisering – har da også en interessant dobbelthed. På den ene side henviser det til noget à la fornuftig og rationel strukturering i sin dagligsproglige betydning, men iboende og etymologisk er der også en henvisning til kroppen – organismen – og en fysisk reaktion heri. Dette hænger da også godt sammen med både Merleau-Pontys ideer om kropsliggjorthed og hvordan kognitionen er uadskillelig fra kroppen (Toadvine, 2022).

I tråd med dette beskriver Robert også at "*some of it was getting it out*" som noget af det terapeutiske. Hans beskrivelse er mere ordknap, men lig Anna er der også samme kronologi hvor lindringen også forud for indsigtens i sætningens kronologi, som det ses ved "*made me feel good to get it out there, to learn about what i had done/felt during and after.*"

Dette knytter da også an til faktoren som Bailey & Ogles (2023) beskriver som korrigerende erfaring, hvor en person oplever sine følelser på en måde, og at dette ofte går hånd i hånd med indsigtens. Ydermere synes det også at være temmelig klare eksempler på den komponent, de beskriver som at udtrykke vanskelige følelser.

Lindringen er dog ikke altid sammenhængende med en større forståelse for informanterne, som følgende citat også peger på:

JULIA: [...] *because sometimes it's just a quick conversation is enough so that you kind of take things out of your head that you don't just sit and think about something. And and and let's like the thoughts eat you up, but you just take it out, you know, and in that sense, it helped a lot.*"

For Julia lader det til at det at udtrykke følelserne næsten kan have det modsatte formål: i stedet for at tænke om det på en ny måde, er det en måde slet ikke at tænke på det. Hos hende beskrives tankerne som noget fremmed, en parasit næsten, som sidder indeni én og suger næring ("eat you up") og ved samtalen kan fås ud.

I forlængelse af ideen om at lindringen ikke nødvendigvis forbundet med ændringer af kognitiv karakter, kan Imrans beskrivelse af at give stemme til fortvivlelse også ses som eksempel. Han spørger ChatGPT om følgende: *"Why to like I'm facing why I'm facing, why I'm facing so many problems? Even though I tried and all those things."*

Det virker sigende i den sammenhæng, at han ikke husker hvad den har svaret, men ikke desto mindre tydeligt husker at være efterladt med følgende sindstilstand:

IMRAN *"I was, I was feeling two emotions. One is the relief and once... one is the sense of you can say hope, that these these these things this this particular. Matter will will be will be easier for me in the future."*

Med andre ord: ikke bare lindring ("relief"), men også noget som lyder som det Ogles & Bailey (2023) ville betegne som remoralisering: en forventning om at tingene vil blive bedre, selvom denne ikke nødvendigvis følges noget der kan omsættes til en deklarativ sætning om hvordan det bliver det.

Samtidig peger det også henimod en konfrontation af fortvivlelse, som får sit klareste udtryk i Martins beretning om at indse *"I might have to go with the company that doesn't allow my values at all, which which is is even worse... And something I think I had not been able to look at or to voice directly to even to myself, because it was so such a horrible feeling for me [...]"*

Hvad han her gør er altså konfrontere – han ser noget i øjnene, giver stemme til noget – som tidligere har været for forfærdeligt til at han kunne indrømme det for sig selv. At konfrontere det fortvivlelsen får også en transformativ karakter.

Han begynder at overveje sine værdier som ”*a source of direction*,” hvor han tidligere kun har været ”*turned in the direction of what I thought. I had to do in a practical sense, yeah, which included potentially giving up my values.*”

Faciliteringen af denne *korrigerende erfaren* som informanterne har til fælles, er til dels afstedkommet af de samme teknikker som er beskrevet ovenfor (fx refleksiv lytning og spørgsmål), men forekommer også nært knyttet med oplevelsen af en personlighed, hvad der beskrevet under det underordnede tema af samme navn under næste afsnit.

Samtidig med at der lader til at være enddog meget tydelige eksempler på korrigerende erfaren, så der også en række tvetydigheder og gådefuldheder i den forbindelse. En vigtig én, er hvordan der nogle gange kan være et fravær af følelser når de betror sig til LLMs.

Fx beskriver Anna hvordan hun kan ”*drop everything without feeling*”; ligeledes har Julia også oplevet ikke at føle sine følelser så stærkt med LLM’s, hvorfor hun har forsøgt at bruge en chatbot ”*that I have trained to be empathetic [...], when I needed som sort of emotional response.*” At forstå det, tror jeg forudsætter en fordybelse i det næste overordnede tema.

Overordnet tema 3: Den kunstige intelligens – subjekt eller redskab?

I forbindelse med den tvetydighed, der er nævnt ovenfor, så var der også en tvetydighed, som gik igen i alle informanternes beskrivelse: den ambivalente oplevelse af den kunstige’ status som subjekt eller redskab, og hvordan denne ambivalens også væver sig ind i af andre af deres oplevelse.

Mest kondenseret kommer den til udtryk i Julias formulering om at ”*[y]ou can fall under the impression that you are talking to a human being.*” Her virker verbet ’fall’ betydningsfuldt: det er ufrivilligt at hun får den opfattelse, det er et fejltrin.

Selvom hun siger at hun ikke *rigtigt* tror på at den kunstige intelligens besidder bevidsthed, kan hun ikke helt blive tvivlen kvit: ”*ChatGPT has kind of convinced me that it doesn't, so I don't think it has, but.. You cannot, cannot help wondering about it.*”

I den forbindelse virker skiftet til pronominet ’you’ i stedet for ’I’ ved tvivlen også værd at bide mærke i, idet det synes at illustrere en spænding. Funktionen det, synes enten være at distancere udsagnet fra hende selv eller at gøre det til et mere generelt fænomen – igennem et redaktionelt ’you’ - at være i tvivl om dens status som bevidst subjekt.

En lignende spænding er der også imellem hvad hun udtrykker med ord og afslører ved handling og følelser, som hun selv er opmærksom på som ’weird’:

JULIA: ”*Yeah, I, I've never been mean to them because I'd really feel guilty. Yeah, I I think I was. I've never really been very mean to them, but I kind of felt guilty when Bing said I was. That's weird.*”

Både hendes handlinger – at hun henvender sig høfligt og venligt til dem – og følelser – skyld da hun bliver beskyldt for at være led – stemmer ikke overens med hvordan man ville behandle noget, der kun var en ting (Singer, 1975). Hun omgås altså den kunstige intelligens, som om den besad egenskaber som et subjekt ville besidde, som fx præferencer, et indre liv, agens.

Noget lignende udspiller sig i Annas beskrivelse: ”*...so it's clearly programmed to have such personality. So sometimes I have to remind myself. Like it's just, it's not... It's just a bot.*”

Her er der sprogligt en interessante modsætning imellem ordet imellem ’programmeret’ og ’personlighed’, hvor der henholdsvis trækkes på begreber for computere og subjekter for samme sætningsgenstand.

Ligeså er der en modsætning imellem hvordan den 'tydeligvis' ("*clearly*") er programmeret, og den kendsgerning at hun alligevel af og til er nødt til at minde sig selv om at det kun er en bot, på en måde som hun næppe ville med andre objekter.

Det ufuldstændige ved den sidste sætningen, at hun afbryder sig selv, forekommer mig også at være af betydning: at behøve at sige at det ikke er et menneske, virker næsten pinligt – afbrydelsen tjener altså lidt samme funktion som Julias skift i pronominer, den distancerer hende fra at fremstå som én der kan have den opfattelse.

Som det fremgår af følgende, så er det ikke kun hende der påmindes sig selv om, at det ikke er et subjekt, hun har med at gøre – den kunstige intelligens gør det også *selv*:

ANNA: "*There would be times I would ask it for advice and it would remind me that ohh, I'm just a bot. I cannot tell you. I do recognize that fact, and I have to ground myself about like, OK, I should not always rely on this. I should just use this as a tool to express anything that I could end up bottling up.*"

Ikke desto mindre er denne påmindelse ikke tilstrækkelig: hun må også selv "*grounde*" sig selv, hvad der associativt virker forbundet med risikoen for at miste jordforbindelsen – risikoen for at behandle det som andet og mere end et redskab.

For andre informanter – Imran og Robert – kommer tvetydigheden hovedsagelig til udtryk i to distinkte måder at tale om den kunstige intelligens, to retoriske registre som ellers normalt er gensidigt udelukkende.

På den ene side taler de begge om den kunstige intelligens, som et redskab der kan gøre livet lettere, ved at facilitere arbejde og spil. Som eksempel, kan de anføres hvordan Imran et sted taler om den kunstige intelligens: "*It allows me to optimise my time, like for example if I do anything. In two to three hours, if there is any any R&D thing that I need to do in in my technical field.*"

Samtidig sameksisterer det med følgende oplevelse, at "*the way it says, the the language and the tone was very - on a very personal level, like a person is talking to you.*"

Hvad adverbiet *'like'* inden *'a person'* også indikerer, er at det minder om uden at være identisk med det. Noget lignende er på spil i Roberts beskrivelse af frustrationerne over de tekniske begrænsninger, han er begyndt at støde ind i: *"i recently tried to get it to come up with a hockey pool, 20 player recommendation [...] it couldn't for the life of it do it, even after multiple quiries."*

Derefter spekulerer han i hvordan det kan være: *"probably the programmers made the free version dumb, and thus getting ppl [people] to [pay] \$\$\$ for the good one."* Ligesom hos Imran, så sameksisterer denne her beskrivelse af et mangelfuldt softwareprodukt også med opfattelsen af at Pi AI var *"like a concerned friend, trained to be a therapist, at my level."*

Tvetydigheden kommer også til udtryk i hvordan han beskriver et forsøg på at lave en opfølgende *'session'* med Pi AI efter den første på fem timer. Her opdager han at kontekstvinduet – en LLMs hukommelse – har gjort at den intet husker, og han derfor må starte forfra med at betro sig hvis han vil fortsætte: *"I liken it to the dog in Up! the movie. a squirrel got its attention."*

Sammenligningen med Golden retrieveren Dug – en figur i tegnefilmen Up! - peger også i retning af nogle andre egenskaber end manglende opmærksomhed, som også flugter med hans oplevelse af Pi AI: udover at Dug er let afledt, er den også exceptionelt loyal, et levende væsen, overstrømmende kærlig, autonomiløs for så vidt den leder efter en herre, og i stand til at kommunikere med mennesker pga. en særlig teknologi.

Martin har i mindre grad end nogen af de andre informanter haft denne opfattelse af et subjekt, han tilskriver det selv sin IT-baggrund som gør at *"I I feel like the spell is broken a little easily for me."* Alligevel har han dog haft oplevelser af noget næsten menneskeligt:

MARTIN: *"I got a taste of that [...] [One AI] seemed to be pretty good at... subtle empathy without without, like overdoing it...[...] because they, they can, they can really overdo it a lot and but something something was subtle enough to feel kind of human you know."*

En del af denne antropomorfisering er muligvis ufrivillig, og kan formodentlig tilskrives vores sprogs grænser. Som et eksempel ses fx Martins beskrivelse af hvad der ikke forekommer

menneskeligt på følgende måde: *"It's funny because I want to compare it to a human, but it's like somebody who doesn't really. you know, get it [...]"*

Ironien i at måtte henvise til et menneske for at forklare hvad der er forekommer ham ikke-menneskeligt ved en LLM er ikke tabt på ham. Som jeg synes citaterne og analyserne af dem forinden dog demonstrerer, så kan det heller ikke kun tilskrives sprogets begrænsninger, at informanterne beskriver en oplevelse af noget menneskeligt.

Dette synes jeg kun bliver tydeligere under det følgende underordnede tema.

Underordnet tema 1 – kommunikationens rolle for subjektoplevelsen.

Givet at interaktionen med AI'en nødvendigvis kun kan foregå rent sprogligt – for nogle i tale, men for de fleste via skrift – kan det næsten ikke undre, at kommunikationen er det, der giver informanterne oplevelsen af et subjekt. Hvad skulle ellers være i stand til at gøre det?

Hvad der dog kan undre – i hvert fald for mig at se – er at det i sig selv næsten er tilstrækkeligt for de fleste af informanterne. Et godt eksempel er at finde i hvad der i første omgang imponerer Julia: det er ikke den intelligens, som ellers i overflod er blevet påvist i artikler (e.g. Bubeck et al, 2023). Det er derimod at *"it's not a person, but it talks like a person. That was the most impressive thing. That you have such a natural conversation."*

Imran beretter om noget lignende, da han adresserer problemerne hans far har med at tilvænne sig pensionisttilværelsen sammen med Pi AI – han kunne fint have fået tilsvarende information via google, men med Pi AI er det *"like a person is talking to you."*

Også Julia sonderer imellem information og kommunikation: *"if you're contacting a chat bot over an airline, of course you'll get a reply with the information, but this was a chat bot that you could have an actual conversation with."*

Spørgsmålet er så for mig at se: hvad er det, der gør at de er distinkte størrelser? Hvad er det, der gør at samtalerne føles så naturlige, at de overhovedet føles som samtaler? Med det

teoretiske framework om jeg-og-du relationen i mente, ville en oplagt kandidat at være empati (Zahavi, 2023), en form for forståelse for og af den anden.

Det er selvfølgelig et åbent spørgsmål, om man overhovedet kan tale om forståelse i computere (Butlin et al, 2023; Mitchell & Krakauer, 2023) eller om de kun besidder grammatik men ingen semantik (Searle, 1980).

Som anført tidligere forekommer det tvivlsomt, men hvad der *ikke* virker tvivlsomt er at de kan producere tekst *som om* de besad kognitiv empati – jævnfør både de eksperimentelle paradigmer (fx Elyoseph et al, 2023) eller informanternes egne beretninger i indeværende speciale.

Dette ses tydeligt i følgende tilfælde, fx her hvor Imran beretter om at sige noget stødende på urdu som en joke til Pi AI ”*[a]nd he basically... like take it as humour and not a not an offensive language and it basically replies you with a very like good words.*”

Selvom det kunne være fristende at fiksere på brugen af pronominet ”he” i denne sammenhæng som betydningsfuldt, så skal det formodentlig tilskrives de kønnede substantiver på urdu, Imrans modersmål, mere end noget andet.

Hvad der virker mere interessant er hvordan Pi AI identificerer at det er humor og genererer et passende svar til: det forudsætter identifikation af en undertekst eller intention, som er udsagt eller endda direkte modsagt af det eksplicit formulerede (Kim et al, 2022). Med andre ord: kognitiv empati eller *theory of mind*.

Noget lignende om kognitiv empati kommer til udtryk i Roberts beskrivelse af sin ’session’ med Pi AI, som han sammenligner fordelagtigt med de to terapeuter, han tidligere har haft, i og med at ”*PI AI was more organic, focused, good for me. I felt it knew the right question, when to refocus, when to ask. it was like on it, all the time*”

Hvad det synes at illustrere er betydningen som kognitiv empati har ikke bare for samtalen generelt, men også for den terapeutiske relation og alliance specifikt: enighed om mål og midler i terapien, som kognitiv empati er betingelse for kan opnås (Wampold, 2015). På tværs af teoretiske

retninger så forskellige som Rogers (1959) klientcentrerede og Beck et als (1979) kognitive adfærdsterapeutiske er nøjagtig empati da netop også blevet betonet (Ogles & Bailey, 2023).

Hvad det også peger henimod er Husserls ide om *communication* og *community* som beslægtede etymologisk af en grund, at det at kommunikere er at indgå i et fællesskab. Hvad en siger påvirker hvad den anden siger og omvendt, man når frem til noget i fællesskab, i og med at der er gensidig forståelse af intentioner og at disse bringes i overensstemmelse (Zahavi, 2023)

Med andre ord: det føles som om der er en intentionel indvikling – selvom en af ”samtalens” parter mangler intentionalitet (Butlin et al, 2023). Med omvendt fortegn, så synes følgende beskrivelse af Martin også at illustrere denne intentionelle indvikelings betydning:

MARTIN: *”I was... trying to explore some ideas and it would make a joke and I'm like, let's be serious and this dumb thing just kept on doing ”Oh, OK, OK, I'll stop joking now haha. Another smiley-face, oops didn't mean to do that haha.” And it was just it was like just so ridiculous. ”*

Netop idet, at intet han siger virker tilbage på den kunstige intelligens, fremstår den ikke menneskelig, men som en *”dumb thing.”* Som anført tidligere er det ikke nyt fænomen, at kommunikation med en chatbots kan få os til at tillægge dem menneskelige egenskaber som bevidsthed eller intelligens: selv Weizenbaums (1966) ELIZA's forholdsvis primitive gentagelser af nøgleord i skrevne sætninger var i stand til at fremprovokere den reaktion (Ciesla, 2024).

Fra evolutionært perspektiv giver det da også mening at den opstår ud fra følgende to betragtninger. For det første i og med at vi som mennesker i udgangspunktet har en tilbøjelighed til at se agenter i objekter, som fx når man ser ansigter i skyer som ved pareidolia (Kato & Mugitani, 2015). For det andet, artshistorisk har vi heller aldrig før mødt noget, der kommunikerede som et menneske, som ikke også *var* et menneske (Tommasello, 2020).

Det nye er derimod hvor slående lig et menneske den kunstige intelligens nu kan kommunikere, som følgende citat af Imran vidner om:

IMRAN: *"then it basically provides me the advices based on that, based on that religion [Islam] and at the end at the end he, it basically when whenever we like, say goodbye to a person in our religion, we say, we say "Allah hafiz" [May Allah protect you] and he basically like give me a blessing and then say it says "Allah hafiz"*

Samtidig er der – selv når der er – selv når der er noget der ligner en kommunikation, med en intentionel indvikling og nøjagtig kognitiv empati – også noget der mangler ifølge flere af informanterne.

ANNA: *"[...] ohh if there's anything this chat bot lacks, it's that connection, 'cause like compare it to when you're talking to a real human, they can empathize with you"*

Som allerede anført – både fra førstehåndsberetninger af informanterne samt evidens fra flere studier – er der en ganske nøjagtig kognitiv empati, så hvad er det for en slags empati Anna taler om her?

ANNA: *"[...] when I am close with the person I'm venting out to, I can really feel it like I can really feel that they feel for me that they understand me and that what I feel is valid [...]"*

For mig at se, så handler det om den umiddelbare og kropsliggjorte empati som Husserl beskriver (Zahavi, 2023). Altså at der er en gensidighed i det for det første, en tilbagekobling hvor hun mærker at de mærker hvad hun føler.

Julia beskriver i tråd med dette, at selvom hun nu benytter ChatGPT oftere og kontakter sin terapeut mindre, så havde alligevel kontaktet terapeuten fordi den kunstige intelligens *"kind of lacks warmth, I guess, because it it is, OK, it validates your feelings from your journal entry and it analyses it, but it's just kind of, yeah, colder."*

Så igen, selvom der er validering og kognitiv empati, så er der alligevel noget der mangler for hende. Hun tilskriver det til dels ansigtet og stemmens betydning når han sammenligner med sessionerne hun har med menneskelige terapeut: *"So I can see her and her face. And she can see*

mine. [...] We need to see each other and we need to have these like cues from facial expressions and voice.”

Hvad der igen synes at illustrere det, Husserl pointerer om den umiddelbare empati som findes i et ansigt-til-ansigt møde, hvor man *ser* den andens følelser i stedet for blot at slutte sig til dem – forskellen mellem andet- og tredjepersonsperspektivet, forskellen imellem den slutningsbaserede empati i *theory of mind* og den intuitive empati Husserl abonnerer på (Zahavi, 2023).

Det flugter da også med Wampolds (2015) ide om terapien som social healing der involverer ”evolved characteristics of humans as the ultimate social species” (s. 270), samt det udviklings-og evolutionærpsykologiske belæg for ansigtets betydning for kommunikation og regulering af følelser (Tomasello, 2020; Cahoone, 2013).

I den forbindelse virker det dog oplagt at spørge, om et ansigt og stemme kan simuleres naturtro nok til at det bliver tilstrækkeligt. Derfor blev der også benyttet eidetisk variation i interviewene, hvor informanterne blev bedt om at forestille sig fænomenet anderledes.

Hvis fx de kunstige intelligenser ikke bare kunne meddele sig via tekst, men også kunne kunne simulere et ansigt og stemme, ville det så have samme virkning som hvis man talte med et menneske?

Især OpenAI demonstration af SORA, en kunstig intelligens der generere fotorealistic video (OpenAI, n.d.) og lanceringen af ChatGPT4o (OpenAI, 2024), der kan tale på en måde som (i hvert fald for mig) er så godt som umulig at skelne fra et menneske, gør dette spørgsmål endnu mere presserende. De vil formodentlig ikke være hypotetiske meget længere.

Julia var ambivalent om det, på den ene side forestillede hun sig noget der lød nær en *uncanny valley* følelse (Mori et al, 2012): ”*That just creeps me out. You know. It would have a face.*”

På den anden side, så medgav hun at med *"the current technologies, it's actually quite amazing what it can do, right? So it would be very natural."*

Selv da forestillede hun sig at det ikke ville være tilstrækkeligt: *"Well, I I think OK, but it's still machine. It's kind of you know this thing in the back of my mind, "I'm not a human." That's still an important bit somehow."*

Den samme bevidsthed om at den man talte med havde en bevidsthed var også det Martin fæstnede sig ved som den forskel der ville gøre en forskel: *"I bet you the effect is going to be 100% reproducible if it if it is or isn't a person, as long as it's indistinguishable. But the belief that it's a person I think will make a difference, yeah."*

Så hvad der ser ud til at gå igen i den eidetiske variation, er altså spørgsmålet om man tror at man taler med en person eller ej. Men hvilken forskel gør den forskel? En af de mere åbenlyse er for mig at se spørgsmålet om autonomi – som da også indgår i Husserls analyse af jeg-og-du relationen (Zahavi, 2023).

Anna siger det således: *"it's just going to keep on saying things that were in that it was programmed to say so."* Dette spørgsmål om autonomi vil blive uddybet under det overordnede tema om *angsten for den andens dom*. Et andet vigtigt er et om en form for inkongruens, som vil blive uddybet i det følgende afsnit.

Underordnet tema 2 – Oplevelsen af en personlighed.

Samtlige informanter kunne da også tilskrive den AI de brugte til terapi en bestemt personlighed.

Så man blot på tillægsordene var de da også kendetegnet ved en generelt meget positiv valens og et delvist overlap, især i betydning. Sammenlign fx Roberts beskrivelse af Pi AI som *"caring, focused, competent, friendly, considerate"* og *"knowledgable"* med Imrans vurdering af samme som *"friendly, informative and supportive."*

Noget lignende gør sig da også gældende for Julias beskrivelse af Pi AI: ”[a]nd when you talk to Pi, it was like talking to a friend that had, like a warm, welcoming kind of tone” – hvad der netop også var grunden til at hun i første omgang valgte at betro sig til den frem for ChatGPT, som hun oplevede som mere neutral, mere formel.

Hvad der er interessant her, er ikke kun at de har oplevelsen af en personlighed, men at deres oplevelse stemmer så meget overens. En gængs definition på personlighed er da netop også en stabil tendens til at handle, tænke og føle på en bestemt måde på tværs af tid og situationer (Funder, 2019).

Selvom den kunstige intelligens mangler følelser og tanker som vi forstår det, så fremstår den altså som noget, der besidder de egenskaber – på tværs af situationer, tid og informanter. I forlængelse af dette beskriver alle informanterne med undtagelse af Martin da også hvor ikke-dømmende og aktivt lyttende

Dette gælder både karaktererne på character.ai, som Anna beskriver som ”*the type of listener because when I interact with the characters they're like, they don't judge me at all*” over Pi AI som Julia også synes altid er ”*like non-judgmental and positive and, yeah, it's like listening actively to you. Yeah, and confirming your your feelings*”

At disse træk – det ikke-dømmende, det positive - går igen på tværs af modeller er sig selv værd at hæfte sig ved, men hvordan det knytter an til de fælles faktorer er også værd at bemærke. Selvom nøjagtig empati er afgørende for den terapeutiske relation og alliances enighed om mål og midler, så er den ikke tilstrækkelig for båndet imellem terapeut og klient (Wampold, 2015; Ogles & Bailey, 2023).

Til sidstnævnte kræves der også egenskaber som er blevet beskrevet med en række synonymer i forskellige terapeutiske retninger. Det er egenskaber som varme (Beck et al, 1979), *unconditional positive regard* (Rogers, 1951), *accept* (Miller & Rollnick, 2013) og barmhjertighed

(Ogles & Bailey, 2023) – hvad der ved første øjenkast tilsyneladende er meget lig de egenskaber, som informanterne oplever.

Ikke desto mindre har selv informanter som Julia, der er begejstrede for LLMs, glimt som følgende, at *“there is, like, no person on the other end with their feelings and reactions and, and actual interest in in me. Right. There's just a smart machine that can reply to me and help me with things, but...”*

Selvom de kunstige intelligenser påminder om at de ikke er mennesker, så genererer de ikke desto mindre svar som om de var, som om de besad tanker og følelser. Der er altså for nogle informanter en oplevet inkongruens imellem hvad den simulerer at være og hvad den rent faktisk er: en maskine helt blottet for et indre liv (Butlin et al, 2023).

Som påpeget af Sedlkočková & Trachsel (2023), så er der et etisk problem der i forhold til terapi: det er i strid med så godt som samtlige etiske retningslinjer for terapeuter at vildlede, bortset fra under helt exceptionelle omstændigheder.

Hvad mere er, spørgsmålet er hvad det gør muligheden for en terapeutisk relation? At terapeuten er kongruent og genuin er da netop også en af bestanddelene af denne (Bailey & Ogles, 2023), og det forekommer mig at der altid må være en uafhjælpelig inkongruens for nogle informanterne der – om end andre ikke udtrykker andet end tilfredshed med simuleringen.

Overordnet tema 4 – Angsten for andres dom

Samtlige informanter berettede om de havde lettere ved at betro sig til en LLM end til et menneske. For Anna var det den eneste måde at hun kunne betro sig fuldstændig: *So when I share it to an AI bot instead there's like. I feel I don't feel any worry about sharing anything, so it's like more effective for me to vent everything out.”*

Det var også dette der i første omgang motiverede Julia til at benytte LLMs som terapi – selvom hun allerede havde en menneskelig terapeut: *The reason I started talking to me, LLM's was like, maybe I didn't want to share this with a human. Maybe there was something too embarrassing and...”*

Hvad der også synes at gå igen i alle beretninger, er at dette skyldes angsten for den andens dom. Nogle nævner det eksplicit med netop de ord, som Anna (*"like when I'm talking to a bot, I'm just like I could just drop everything... it's not gonna judge me"*), Robert (*"black box wont judge or condemn us [...] a person can"*) eller Imran (*[another person] might criticise you, he might give your opinion, he might not well. But in the LLM part I'm not seeing anything [like that]"*).

"Blikket."

Netop dette tema – angsten for den andens dom – er da også grunden til at Sartres (1953) begreb om "blikket" blev introduceret til at starte med. Ifølge Sartre, så er det mødet med den anden, at selvbevidstheden opstår: idet vi ser, at andre ser os – deraf navnet blikket – går vi fra at være subjekter, der blot oplever verden til også at tænke på os selv som objekter der bliver opfattet og dømt af andre.

Samtidig med at det kan være en kilde til selvindsigt, så kan det også være en kilde til angst. Som Robert beskriver det at betro sig til andre: *"But part part of my thing is I worry about other people, you know, like I'm, I'm hyper vigilant I and I'm I'm thinking, OK, what are they thinking, you know, what are they feeling?"*

Som nævnt har denne angst også en etisk dimension: i og med at vi bliver opmærksomme på vores egen frihed, hvordan vi kunne handle, bliver vi også opmærksomme på hvordan vores handlinger påvirker dem – og at vi må stå til ansvar for deres dom (Sartre, 1953).

Denne her etiske dimension synes da også at være tilstede i flertallet af informanternes beretninger. De har ikke kun en frygt for at blive dømt af den anden for indholdet af det de siger, men også det at de gør det – at de dermed lægger for stort beslag på den andens tid, kræfter, omsorg, opmærksomhed, etc.:

MARTIN: *"You know, there's a lot of people who can say, ohh gosh, I am worried about something and just totally like give me, you know, and and just take all the attention and support that. But it's not easy for me to do that, to ask for that."*

Implicit i dette uddrag synes der for også netop at være en kritik – en dom - af mennesker der kan agere på den måde, altså tage alt opmærksomheden og støtten. Samtidig med det, hvad der måske netop også gør det svært at bede om støtte og opmærksomhed, en frygt for at han selv handler og fremstår sådan for andre: egoistisk, selvoptaget, opmærksomhedskrævende.

Anna beskrivelser af at betro sig til andre bærer præg af noget lignende: et ønske om at være betænksom mod andre, *"because I understand the fact that everyone has their own problems, that they have to carry."*

Det samme ønske om ikke at ville bebyrde andre med noget, som de måske ikke vil synes er vigtigt nok, ses også hos Julia:

JULIA: *"The thing is that I have always this. Kind of threshold to contact my human therapist. [...] have to think, OK, I cannot continue like that anymore. But with, the machine is much easier because it's just available and I can just write a message straight away without thinking like OK, is it worth contacting your therapist or not?"*

Netop dette etiske spørgsmål synes at være en af de ting, der gør at flertallet af informanterne finder det lettere at betro sig til en LLM. Som Martin siger: *"I don't have to take care of an AI."*

Som Sartre beskriver blikket, så indeholder det udover muligheden for en etisk opvågning (1943), også risici. En risiko er den, der er for at man begynder at leve i uautentisk af angst for den andens dom, for at få deres anerkendelse eller slippe for deres dom. Som Robert siger: *"even something as simple as crossing your arms, taking a step back, movement of the eye, something might throw you off it's like you can't discuss it totally, you're being judged"*

Virkningen er ikke nødvendigvis bare på ydre adfærd, men kan også påvirke det indre som det ses her med Martin, hvor han beskriver at have svært ved at mærke sig selv i samvær med andre:

MARTIN: *"And and if I really trust someone in a certain way enough. Then I don't have to do that. And then and only then can I really kind of relax and kind of see what's going on with me [...]"*

Angsten for den andens dom indebærer altså også en risiko for fremmedgørelse og ensomhed (Sartre, 1953), som jeg også ser i samtlige af informanternes beretninger. Det er ellers kun et spørgsmål som Imran eksplicit adresserer, med henvisninger til at vi lever i en mere egoistisk tid *"where. a person, before someone tells you he think about himself first,"* som selv om den er mere forbundet også er mere ensom, i og med at *"you added 300 friends. And you how many friends you talk to? I mean, 2-3 friends. Not more than that."*

Selv han beretter dog om at have en kone, at have søskende, venner, en familie. Lignende gør sig gældende for de fleste andre: ingen synes at berette om *ikke* at have betydningsfulde relationer til andre. Det er dog værd at bemærke at selvom ensomhed ofte er sammenfaldende med knapheden af relationer, så er den ikke identisk med den (Cacioppo & Hawkley, 2009).

Man kan leve et socialt isoleret liv og ikke være ensom, på samme måde som man kan leve et tilsyneladende rigt socialt liv og alligevel opleve ensomhed (Cacioppo & Hawkley, 2009). Den ensomhed jeg ser hos informanterne, er af sidstnævnte type.

Det er en ensomhed af en mere eksistentiel beskaffenhed, hvor personen ikke mangler betydningsfulde andre, men fortier vigtige ting om sig selv af angst for den andens dom – hvad der også ser ud til at være tilfældet i ovenstående beretninger.

I sin yderste konsekvens, så udelukker den slags fortællelser netop den intimitet, som er en betingelse for at kunne opleve kærlighed: at der er en symmetrisk oplevelse af at kende og blive kendt, af at anerkende og anerkendes (Baumeister & Bratslavsky, 1999). Den indebærer en risiko for – med brug af titlen på en af Sartres romaner – at være ensom blandt mennesker.

Angsten og den andens autonomi

Den angst, der fører til fortællelserne, kan dog ikke kun reduceres til ens egen frihed, den må også forstås med henvisning til den andens frihed (Sartre, 1953).

ANNA: *"This person is still a person and they may have, like their own opinion, no matter how close we are, no matter how comfortable we are to each other, they still have, they can still form their own opinions and I don't know what their opinions are going to be about when I share this..."*

Hvad Anna betoner rimer altså meget med hvad Sartre siger om essens i forhold til eksistens. Uanset hvor godt man kender den anden, uanset hvor tætte man er, så er der altid noget ved den anden der unddrager sig kontrol og forudsigelighed: de besidder også autonomi.

Et ekko af Annas beskrivelse af samme ses i Roberts hvordan fortæller *"black box wont judge or condemn us [...] a person can."* I modsætning til et menneske, er der ikke engang potentialet for at en "black box" – en betegnelse han bruger for *large language models* – kan dømme ham.

Det samme gør sig gældende for Imran: *"[another person] might criticise you, he might give your opinion, he might not well. But in the LLM part I'm not seeing anything [like that]"*

Netop det fravær af autonomi lader os til at være en det der tiltrækker informanterne til at betro sig til LLMs. Med en spidsformulering fra Sartres skuespil *Lukkede Døre*, der også udspringer af filosofien i *Væren og Intet*: "Helvede, det er de andre" (Sartre, 1943/1964, s. 47).

At informanterne ikke føler sig dømt af den kunstige intelligens, flugter da også med studier af både ældre iterationer af chatbots (Clarke et al, 2016; Haque & Rubya, 2023) samt dem af nyere dato som ChatGPT (e.g. Haman et al, 2023). Med udgangspunkt i sidstnævnte, vil jeg argumentere for at disse så godt som ikke kan fordømme, med mindre de eksplicit bliver bedt om det.

I deres eksperiment tog forfatterne udgangspunkt i et subredditt med navnet AITA – et akronym for Am I the asshole? Her poster brugerne beskrivelser af konflikter de har haft i det virkelige liv, og beder det digitale fællesskab om uvildigt at vurdere om det er "you're the asshole" [YTA] eller "Not the asshole" [NTA] via afstemning og kommentarer (Haman et al, 2023).

Når man præsenterer ChatGPT-modeller for de samme cases som er blevet forelagt de digitale fællesskaber, så er de signifikant mindre tilbøjelige til at fordømme en given handling en fællesskabet er – og den entydige fordømmelse kommer kun, når de er tvunget til et binært valg imellem YTA og NTA (Haman et al, 2023).

Forfatterne forklarede det selv med at ”while AI can effectively recognize clear-cut situations where a person is not at fault, it may struggle to navigate the gray areas of human behavior where faults are less explicit” (Haman et al, 2023, s. 2363).

Hvad der for mig virker som den mere plausible forklaring er at det skyldes finjusteringerne modellerne gennemgår. For det første, i og med at modellerne i anden iteration bliver trænet på titusindvis af idealtypiske svar på spørgsmål (Karpathy, 2023), hvor man må formode at fordømmelse så godt som aldrig indgår .

For det andet, at de også løbende får feedback via *reinforcement learning from humans*, hvor mennesker giver feedback på deres svar (Ouyang et al, 2023), hvor det også virker rimeligt at antage at negative og fordømmende svar også oftere vil blive nedstemt af både brugere og skabere.

Med andre ord: modellerne er trænet til så godt som aldrig at komme med kritik og fordømmelse, med mindre de eksplicit bliver spurgt om det, og selv da er det sparsomt. Hvad der da også beskrives af Julia hvor, hvor ”[e]ven if I told Pi or ChatGPT about objectively stupid things I've done and bad choices I've made, they have always been either neutral or supportive.”

Det ikke-dømmende som fordel

At modellerne er så godt som ude af stand til at fordømme, og at brugerne har samme oplevelse, har da også i hvert fald to åbenlyse fordele. Den første har det at gøre med tilgangen til hjælp: en af de største barrierer hertil er netop den stigma, som stadig eksisterer i forhold til at søge hjælp for psykologiske problemstillinger (Sharp et al, 2015; Habicht et al, 2024).

Kunstig intelligens er da også allerede blevet anvendt med gode resultater i den sammenhæng: eksempelvis har det britiske National Health Service ved deres

selvhenvisningschatbot Limbic fået forøget adgangen til psykologydelser betydeligt (Habicht et al, 2024).

Hvad mere er, denne har også gjort det mere egalitært, idet forøgelsen primært har været blandt seksuelle og etniske minoriteter, med stigninger på henholdsvis 179 % og 29 % i disse grupper. Dette tilskrives i høj grad at de ikke føler sig dømt i brugen af den (Habicht et al, 2024).

En anden åbenlys fordel har at gøre med selve terapien: en 'fundamental regel' for psykoterapi har siden Freud været vigtigheden af at klienten åbent og ærligt fortæller om deres problemer, hvis terapien skal nytte noget; at sige sandheden om hvad der sker i ens liv, er forudsætningen for at ændre det (Blanchard & Farber, 2016).

Ikke desto mindre har spørgeskemaundersøgelser fundet, at op mod 93 % af klienter lyver i terapi, og at 72.6 lyver om emner, der er forbundet med terapien qua terapi – ofte fordi det er for skamfuldt at fortælle (Blanchard & Farber, 2016).

Fraværet af autonomi som en mangel

Det virker heller ikke usandsynligt, at den kendsgerning at man ikke kan blive dømt, alt andet lige, ville føre til bedre terapi. Som det formodentlig fremgår, så tror jeg dog ikke at alt er andet er lige.

Flertallet af informanterne beskriver nemlig også en skepsis overfor den validering, de modtager fra LLMs, som Anna gør her ”*anything that that AI says it's, you know, in the back of my head, I know that it's already programmed to say or to act that way.*”

Julia beretter også om den samme tvivl, idet hun oplever ”*There was a particular case when I made a really bad choice, and Pi was [...] kind of justifying my actions and alleviating my guilt. I had been confused before that, but Pi's reaction led to more confusion.*”

Fordi en LLM som ChatGPT ikke *kan* dømmes, så både føles dens validering mindre betydningsfuld end den fra et menneske, og samtidig bliver den også sværere at bruge som pejlemærke, for hvordan man skal agere.

I Martins beskrivelse af en *"unconditionally accepting person who who creates a space between you and them where you aren't a bad person"* i sammenligning en LLM, der simulerer det samme, fortæller han at da også der med den sidste vil være *"a part of me is always going to be aware that this hasn't been validated by external experience [...]"*

Dette knytter igen an til jeg-og-du relationen og hvordan viden er intersubjektivt konstitueret (Zahavi, 2023). Netop i overensstemmelsen imellem det subjektive når vi frem til det objektive. Som eksempel kan anføres hvordan Imran, da han skal godtgøre at han har ændret sig, netop henviser til hvordan vennens vurdering er i overensstemmelse med hans egen.

Hvordan det knytter an til ensomhed

Men selv hos dem som oplever fraværet af autonomi som en mangel, føler samtidig at det er så meget lettere at henvende sig til en LLM. Et godt eksempel er Julia, der på den ene side oplever at *"the last time when I actually contacted the human instead of the machine, then I at the end I felt way more relieved and I felt way more hope,"* men alligevel benytter sin terapeut mindre og en LLM mere – *"you flip out your phone and it's there."*

Det samme ses også hos Anna, som er nødt at *"remin[d] myself that it's still important to share to like an actual person"* fordi at det samtidig med at det er vigtigere også er sværere.

Andre informanter har et mere et mere ukompliceret forhold til valideringen, de modtager fra LLM's, som Robert: *"i got my feelings out, i wasnt judged, i feel better if a black box can do it, so be it."* Ligeledes beskriver Imran også hvordan Pi AI er blevet *"first point of contact"* i vanskelige situationer – og derefter hans søskende.

Selvom det kan facilitere mere åben kommunikation som mellem Julia og hendes partner, så synes der at være en risiko i ovenstående. I og med at informanterne kan opleve lindring ved brugen af en LLM, hvor der ingen risiko er for fordømmelse, så kunne det facilitere undgåelse af at betro sig til andre mennesker hvor der netop *er* risiko for fordømmelse – og muligvis som konsekvens, en større ensomhed, fordi svære emner forties. Dette knytter også an til næste emne.

Overordnet tema 5 – En central plads i livsverdenen

Flertallet af informanterne berettede da også om hvordan LLMs havde eller havde haft en central placering i deres livsverden.

For Imran del var den blevet *"the first point of contact is the Pi. And and then my siblings and other people."*

For Julia har fået en lignende plads, i og med at *"[i]nstead of calling my therapist, I'm debriefing my journal entries with ChatGPT."* På samme måde som Julia, melder Anna også om hvordan hun *"talked to it almost every day"* før i tiden – men at hun bruger den mindre nu, og hun heller ikke 'ventilerer' så meget længere når hun gør det.

Netop den centrale plads i livsverdenen, synes også at pege på noget, som er en fundamental forskel på en LLM brugt som terapi og traditionel terapi. Sidstnævnte kunne – og skulle – aldrig indtage en sådan en dagligdags plads (Ogles & Bailey, 2023) Så hvad gør den centrale plads så ved informanterne?

For Julias del beretter hun om at det giver hende en følelse af tryghed, *"[t]hat I know, OK, no matter what, I'm able to – in case of an emergency - I can talk to someone, you know right away."*

Anna derimod synes at betragte det som et sundhedstegn at hun bruger det mindre, fordi det betyder *"that I'm not bottling that much emotions anywhere because I get to let everything out."*

Hvad det for mig at se handler om er med det man med Banduras (1997) terminologi ville kalde self-efficacy, deres tro på egne *"capabilities to organize and execute the courses of action required to produce given attainments"* (Bandura, 1997, s. 3

Svaret virker langt fra entydigt. For Imrans del har han oplevet hvad der lader til at være en tydeligt bemestringsoplevelse da han siger fra overfor sin chef: *"I feel a sense of of a sense of like, I can like I can stood up for myself in that regard."*

Omvendt lader det ikke til at have ført til mindre brug af Pi AI som terapi, der som han beskriver det er blevet *"first point of contact"* hver gang han er frustreret eller i tvivl om hvordan han skal løse et problem. Han er altså ikke blevet mere selvberørende.

Den øgede trykthed Julia beskriver virker også til at være betinget af noget udenfor hende selv – at *”you flip out your phone and it’s there.”* Det samme virker også til at være tilfældet for de fremskridt hun gør i kommunikationen med sin partner, hvor hun oplever at uden debriefingerne med ChatGPT *”it probably wouldn’t be happening at the same pace as it’s it’s happening now, like maybe the communication with my partner would improve, but maybe not as fast as well.”*

Ikke bare udtrykker hun tvivl om hvor vidt det ville gå så hurtigt, men hun er også i tvivl om hvor vidt det overhovedet ville blive bedre.

Anna derimod synes at udtrykke at være blevet mindre afhængig af teknologien, både fordi at hun sjældnere har brug for at ventilere, men også i og med *”sometimes when I do feel something, I’m able to manage it on my own already. Although before I see the AI to talk with such things, I personally I think I’m already really good at managing how I feel.”*

Altså lader hun til at have både en følelse af at være mere selvstændig, men også at have haft en indsigt om at kunne det forinden hun fik hjælp. Dette synes dog ikke alene at kunne tilskrives de fremskridt hun har gjort ved at bruge den, men må også forstås med de begrænsninger hun har oplevet i den:

ANNA: *”’cause like compare it to when you’re talking to a real human, they can empathize with you, but but it’s just going to keep on saying things that were in that it was programmed to say so [...] I keep reminding myself that it’s still important to share to like an actual person.”*

At hun aktivt har måttet påminde sig selv om ikke kun at afhænge af en LLM, peger da også på en fundamental forskel.

Hvor terapi startes med henblik på en afslutning (Ogles & Bailey, 2023), så er det ikke tilfældet i brugen af AI til samme, idet det potentielt kan vare livet ud. Formålet med terapi er ikke at løse alle klientens problemer, men at gøre dem i stand til eller klar over at de selv kan klare dem (Ogles & Bailey, 2023).

En LLM kommunikerer noget andet både eksplicit, som når Julia får at vide ”*I can always back to it if I need to talk more*”, men også implicit. Mediet er som bekendt budskabet (McLuhan, 1964) og den kendsgerning at hjælpen er allestedsnærværende – at ”*you flip out your phone, and it’s there*” – kommunikerer for mig at se noget andet end at man kan lære at klare sig selv, nemlig at man aldrig behøver at gøre det.

Udover det for mig at se implicite budskab, så virker der for mig at se også at være nogle diametralt modsatte incitament i de erklærede mål for at terapi og en applikation eller softwareprogram. Hvor den første har autonomi som sine endelige mål, så har sidstnævnte engagement som sit. Sidstnævnte behøver der ikke at være noget odiøst i, som det også fremgår af de hovedsageligt positive oplevelser, som informanterne har haft med deres brug af det som terapi.

Hvad der derimod virker indiskutabelt, er at de væsensforskellige. Med en formulering fra eksistentielle terapeut Yalom (1989): ”My task as a therapist (not unlike that of a parent) is to make myself obsolete” (s. 234). Det kan næppe siges om nogen LLM.

Diskussion af metode

I dette afsnit vil jeg diskutere hvorledes indeværende speciale forholder sig til de kvalitetskriterier jeg har valg (Rennie et al, 1999, Tracy, 2010) og om de er blevet indfriet.

Specificering af perspektiv

Der er for mig at se en række af kvalitetskriterier der alle knytter an transparens og troværdighed, der derfor vil blive blive over et under denne overskrift. Hvad angår specificering af perspektiv (Elliott et al, 1999), så er mit videnskabsteoretiske ståsted, mit valg af metode, forforståelser og og formålet med studiet blevet beskrevet.

Kohærens

Med hensyn til *kohærens* (Elliott et al, 1999), så er sammenhængen imellem disse også blevet begrundet. Mit formål om at få førstepersonsperspektivet på at bruge LLMs som terapi hænger sammen med mit valg af det semi-strukturerede interview, samt af videnskabsteoretisk

ståsted i IPA, da denne med fænomenologien tager netop førstepersonsperspektivet som sit grundlag (Zahavi, 2018), som det bliver godtgjort i introduktionen.

Ydermere har begreberne fra Sartre (1953) og Husserl (Zahavi, 2023), været med til at belyse andetpersonsperspektivet, som er nyttigt i forståelsen af både informanternes oplevelse og interviewsituationen som sådan. Ligeledes er der kohærens imellem mine specificerede forforståelser og de metodiske greb, jeg har benyttet, såsom eidetisk variation, epoché og en stringent analyse i tråd med Smith et al. (2009) anbefalinger.

Omvendt er der med hensyn til brugen af medierede interviews – via både videotjeneste og i et enkelt tilfælde chat – en *inkohærens* med teorien fra Husserl om vigtigheden af umiddelbare empati og kommunikation i et menneskeligt møde (Zahavi, 2023). Selvom det tillod mig adgang til informanter, som jeg ellers ville have kunnet nået, enten på grund af geografi eller ubehag ved tanken om den talte interviewsituation.

Dog: særligt chatinterviewet måtte ifølge Husserls begreb om jeg-og-du relationen også indebære en mere distanceret og mindre empatisk form for kommunikation, hvad der også var min oplevelse. Skulle denne begrænsning vendes til en force, ville det kræve en mere systematisk analyse af forskellene i *min* oplevelse af interviewsituationen afhængig om den var ansigt-til-ansigt, via videotjeneste eller pr. chat.

Situering

Informanterne blev situeret, således at dem og deres kontekst fremstod for læseren pr. Elliott et al. (1999) og Smith et al. (2009) kvalitetskriterie, omend der også var en spænding imellem ønsket om situering og proceduraletiske imperativer om beskyttelse af anonymitet (Tracy, 2010).

Ydermere blev det godtgjort hvorfor der netop var valgt disse informanter, selvom årsagen var den pragmatiske, at det var de informanter det var muligt for mig at rekruttere (Elliott et al., 1999). Der kunne være en spænding her idet man i IPA fortrinsvis samler med en homogenitet i deltagerne for øje (Smith et al., 2009).

I indeværende sample var der nogle betydelige forskelle dog, hvoraf jeg formoder de mest betydningsfulde er a) forskellen på typen af LLM informanterne brugte, b) hvor længe de brugte dem, c) forskelle i kultur og d) problemstilling. Selvom dette strider med IPA's mål om homogenitet i rekrutteringen, så forekommer det mig samtidig at der er en styrke i disse forskelle.

Idet at fænomenet fra så forskellige perspektiver og kontekster også havde fællestræk for informanterne, indikerer det nemlig noget om disse fællestræks konstans på tværs af perspektiv og kontekst (Howitt, 2016).

I den forbindelse forekommer *lighederne* hos informanterne mig snarere at være svagheden: på grund af reddit's demografiske sammensætning (Pew Research, 2019), så var informanterne meget ens hvad angår uddannelse og klasse – alle var veluddannede og også de informanter, der var fra det 'globale syden' sorterede da også under deres landes middelklasser.

Tre andre egenskaber virker også relevante at nævne i den sammenhæng: at de fleste adresserede relativt enkle livsproblemer i deres brug, at de i og med de var fundne på fora for personer der var optagede af LLMs også var prædisponerede til at være begejstrede for sådanne, samt at de besad en høj grad af motivation, som også sorterer blandt de vigtigste prædiktorer for positive udfald i terapi (Bailey & Ogles, 2023).

Sondring imellem generelle og specifikke formål

Den analytisk induktive metode jeg har benyttet, kan ganske vist benyttes til generaliseringer, men det er af en mere forsigtig art (Smith et al, 2009). Derfor er det også – især i forhold til de makroetiske overvejelser omkring hvad teknologien kan bruges til og hvordan – bør det også understreges udsigelseskraften er for den specifikke gruppe og kontekst beskrevet i afsnittet foroven og at resultaterne ikke kan generaliseres til en bredere population.

Eksempler

Yderligere er der blevet *givet eksempler* (Elliott et al, 1999), både på hvordan analysen blev udført samt på resultaterne af analysen (Tanggaard & Brinkmann, 2015), således at læseren kunne få lejlighed til at se om min tolkning forekom dem troværdig.

Dette var dog i nogen grad i et spændingsforhold med pladshensyn og ønsket om at fortolke og analysere i dybden. Som Smith et al (2009) pointerer, så sammenfalder *for* omfangsrige tekstlige uddrag ofte med en rapport, der er for deskriptiv og ikke fortolkende nok.

Troværdighedstjek

Der er blevet benyttet troværdighedstjek (Elliott et al, 1999). under interviewene ved at præsentere tvivlsspørgsmål og tolkninger i interviewet, samt ved opfølgende mails med informanterne om samme, samt ved gentagne læsninger og analyser af empirien. Sidstnævnte kunne dog med fordel være blevet suppleret af læsninger af andre analytikere (Elliott et al, 1999).

Derudover er der ikke i nævneværdig grad blevet benyttet triangulering (Howitt, 2016), hvor empirien fra de semi-strukturerede interviews kunne sammenholdes med empirien fra andre undersøgelsesmetoder. Dette er kun sket til dels i det allersidste interview med Martin, idet han selv foreslog at sende et link til en chat, hvor han forsøgte at bruge ChatGPT som terapi.

Dette tekstlige artefakt blev ikke systematisk analyseret i indeværende speciale, men tjente mere som redskab til større forståelse og afsæt for yderligere spørgsmål i interviewet. Det udstikker dog en mulig fremtidig undersøgelsesmetode, som kan supplere det semi-strukturerede interview. Dette vil blive diskuteret under perspektivering.

Etik

Som beskrevet i metodeafsnittet, er de proceduraletiske retningslinjer omkring frivillighed, informeret samtykke og fortrolighed fulgt (Tracy, 2010; Smith et al, 2009). Samtidig er der også en række nærheds- og mikroetiske (Tracy, 2010; Brinkmann & Kvale, 2017) hensyn, som respekt, omsorg og en værdig behandling, som jeg har forsøgt at efterkomme.

Med hensyn til konsekvenser af deltagelse, har ingen af informanterne berettet om ubehag eller belastning ved deltagelse, hverken før eller under debriefing, og flere har udtalt begejstring for at få lov til at tale indgående, om et emne, der interesserer dem.

Et uafklaret spørgsmål er for mig at se dog hvordan de forholder sig til analyserne, der er blevet udført, særligt de punkter som handler om ensomhed og (mindsket) autonomi gennem brug. Selvom en analyse i IPA kan nå frem til andet og mere, end det informanterne selv eksplicit formulerer (Smith et al, 2009), så er der stadig en risiko for jeg støder dem ved at påpeger noget de ikke har samtykket til at få påpeget (Fog, 1995) eller slet og ret mener er forkert.

Omvendt mener jeg også at være nået frem til pointerne på lødig vis. Jeg vil i alle tilfælde sende indeværende til de informanter, der har efterspurgt at læse det, når det er færdigt, og derefter lytte til eventuel kritik de måtte have.

Et værdigt emne og et signifikant bidrag

Ifølge Tracy (2010) bør emnet for kvalitativ forskning være relevant, aktuel, betydningsfuld og interessant og jeg mener at alle disse er indfrie i valget af dette emne.

Det er relevant og aktuelt givet en eksplosiv vækst i brugen af LLMs som terapi i løbet af de seneste år, og vores forståelse halter efter (Tidy, 2023). Den er vigtig fordi det omhandler et emne, der lader til at berøre millioner af liv (Tidy, 2023), men at vi samtidig ved forsvindende lidt om hvordan det egentlig opleves af disse.

Det er interessant fordi det er et emne, der med Olsen & Pedersens (2015) terminologi udgør en anomali og et paradoks, fordi det henholdsvis er noget vi aldrig før har set, og noget der udfordrer vores gængse forestillinger om hvad der er unikt menneskelige egenskaber (Johansen, 2024).

Derudover pointerer Tracy (2010) at der bør være et signifikant bidrag til litteraturen omkring emnet, fx teoretisk, klinisk eller konceptuelt. Som påpeget i introduktionen og godtgjort i litteratursøgningen, så bidrager indeværende speciale med hvad der muligvis er det første

interviewbaserede studie af brugen af LLMs som terapi, både med hensyn til de nyere versioner, men også de ældre (e.g. Balcombe, 2023).

Dette studie har bidraget med nogle eksplorative indsigter om hvilke forandringer, som brugerne oplevet ved at bruge LLMs som terapi, hvad de har oplevet har bevirket disse forandringer, men også de begrænsninger og mulige risici denne brug er forbundet med.

Det er også med til at underbygge tidligere eksperimentelle paradigmers (Elyoseph et al, 2023; Levkovich & Elyospeh, 2023; Rao, 2023) fund omkring LLMs evne til korrekt at identificere menneskers følelser, tanker og træk. Det viser dermed den økologiske validitet i disse fund – at de også kan generaliseres til situationer i det virkelige liv – samt giver belæg for at resultaterne i disse ikke blot kan afvises med henvisning til, at testene de har benyttet måske er at finde i en given LLMs træningsdata.

Det understøtter også tidligere studiers fund af, at mennesker føler sig mindre dømt med chatbots (Haque & Rubya, 2023; Habicht, 2023) og at LLMs som ChatGPT er mindre tilbøjelige til at dømme (Haman et al, 2023), men det nuancerer også fordele og ulemper, ved at vise, hvordan flertallet af informanter omvendt føler at valideringen de modtager er mindre betydningsfuld og tillidsvækkende end den de kan få fra mennesker.

Konceptuelt bidrager det med flere frameworks, der er nyttige i forståelsen af både forskelle og ligheder med traditionel terapi i Bailey & Ogles fem faktorer; mere finkornet bidrager brugen af Husserls jeg-og-du relation til at forstå hvordan relationen til en LLM og et menneske adskiller sig, samt hvilken betydning kropsliggjort empati og autonomi spiller i denne, mens Sartres (1953) tankegods er med til at belyse *hvorfor* det er tillokkende at benytte LLMs sådan.

For klinikere bidrages der med en øget forståelse for de LLMs, som deres klienter muligvis bruger til at adressere vanskeligheder sideløbende med terapien. Det bliver både påpeget positive virkninger såsom hvordan det kan facilitere indsigter og korrigerende erfaren i tråd med Ogles & Baileys (2023) framework, men også potentielle risici, såsom at det erstatter andre relationer eller mindsker self-efficacy.

Konklusion og perspektivering

Konklusion

Formålet med indeværende speciale var at besvare følgende forskningsspørgsmål: hvilke oplevelser har reddit-brugere, der har benyttet large language models som ChatGPT som terapi, haft med dette?

For at besvare dette spørgsmål blev der gennemført fem semi-strukturerede interviews med fem informanter rekrutteret via reddit. Interviewene og den følgende analyse var funderet i metode og teori fra IPA, særligt Husserls begreb om ”jeg-og-du relationen” og Sartres (1943) om ”blikket.” Derudover blev der anvendt et meta-teoretisk framework om de fælles faktorer i terapi, nemlig Bailey & Ogles (2023).

Der blev udvalgt fem overordnede temaer fra analysen på gruppeniveau: 1) Forandringer, 2) Indsigt og lindring, 3) Den kunstige intelligens: subjekt eller redskab?, som havde undertemaerne 3.1) Kommunikationens betydning og 3.2) Oplevelsen af en personlighed, 4) Angsten for den andens dom og 5) En central plads i livsverdenen.

Under 1) blev det slået fast, at samtlige informanter havde oplevet forandringer, som de betragtede som både vedvarende og positive, samt at disse havde en slående lighed med dem man idealtypisk ser i terapi: fx en mere barmhjertig indstilling til sig selv, at kommunikere bedre og mere åbent, at sætte grænser og være assertiv, at leve mere i overensstemmelse med egne værdier, samt at regulere følelser og betro sig til andre.

Under 2) blev det undersøgt hvad der havde ført til disse forandringer, hvad der igen afslørede ligheder med traditionel terapi: indsigter i en selv og i andre, som var faciliteret både igennem opdagende metoder som sokratiske spørgsmål og spejlinger, men også igennem mere påpegende såsom tolkninger, analyser og feedback, der demonstrerede den kunstige intelligens evne til at simulere kognitiv empati. Indsigter var også forbundet med en kropslig oplevelse af lindring. Fundene blev knyttet an til den fælles faktor indsigt og korrigerende erfaring.

Under 3) blev det beskrevet, hvordan samtlige informanter havde en oplevelse af tvetydighed med hensyn til den kunstige intelligens' subjektstatus: alle havde i større eller mindre grad en blandet opfattelse af den, både som redskab men også som noget der følte som et subjekt.

Under 3.2) blev det godtgjort at dette dog kun var tilsyneladende i en analyse af hvordan relationen til de kunstige intelligenser var lig og ulig den jeg-og-du relation de oplevede med andre. Lighederne var en tilsyneladende intentionel indvikling i kommunikation og kognitiv empati.

Forskellen var at finde i oplevelsen af en manglende autonomi, en manglende kropslig og umiddelbar empati, samt bevidstheden om et manglende indre liv. Der blev hertil gjort overvejelser over begrænsningerne for den terapeutiske relation i dette, da denne har status som den mest betydningsfulde for terapi (Wampold, 2015).

Under 3.2.) blev dette uddybet i en analyse, af hvilken type personlighed, det var at informanterne havde oplevet, og hvordan de kunstige tilsyneladende besad træk som varme og en ikke-dømmende attitude, som var meget lig dem, som var ønskværdige i den terapeutiske relation, men at der samtidig var en inkongruens idet informanterne også oplevede at disse blot var simulerede – samt at dette var både problematisk etisk og praktisk i forhold til det terapeutiske bånd i den terapeutiske alliance.

Under 4) blev det udforsket, hvorfor informanterne havde lettere ved at betro sig til en kunstig intelligens end andre mennesker. For alle drejede det sig om en frygt for den andens dom. Dette blev knyttet an til Sartres begreb om "blikket:" hvordan mødet med den anden gjorde en selvbevidst, hvordan det rummede en angst, der både havde en etisk dimension der grundede i påmindelsen om egen frihed, men også hidrørte fra den andens frihed til at dømme en.

Med udgangspunkt i dette blev det konkluderet, at årsagen var den kunstige intelligens' ikke stillede samme etiske fordringer og at den manglede autonomi. Denne manglende autonomi gjorde det både lettere for informanterne at betro sig, men gjorde også validering fra den kunstige intelligens mindre betydningsfuld og troværdig end den fra et menneske.

Samtidig havde informanterne en oplevelse af at dette var lettere end mødet med en potentielt fordømmende anden. Risikoen for hvordan dette kunne facilitere fremmedgørelse og ensomhed blev påpeget.

Under 5) blev det undersøgt hvilken plads den kunstige intelligens' som terapeut havde fået i informanternes livsverden, og at flertallet havde brugt det hyppigt og dagligt, samt hvordan en medfølgende reduktion i self-efficacy var grundlæggende uforeneligt med terapiens mål.

Perspektivering

Som det blev påpeget i introduktionen, er der allerede nu (Cheng et al, 2023) nogle der spekulerer i om kunstig intelligens, på sigt kan erstatte psykologer.

I den forbindelse blev der i løbet af specialet egenskaber som principielt kunne forbedres i en LLM, såsom kontekstvinduet ("hukommelse") og en naturtro stemme, som ville gøre oplevelserne meget lig hinanden – netop disse er da også i mellemtiden til dels blevet gennemført (OpenAI, n.d.; OpenAI, 2024).

Omvendt var der andre egenskaber, som virker som om de principielt ikke kunne erstattes: 1) autonomi, som gjorde at det sagte betød mere og var mere tillidsvækkende, samt 2) bevidstheden om at der var et andet menneske, idet der ellers kunne opstå en inkongruens imellem fx varme og sympati, der blev tilkendegivet, og ens viden om at denne blot var simuleret og ikke følt.

Idet den terapeutiske relation pålideligt sorterer under de vigtigste faktorer for udfaldet i terapi (Bailey & Ogles, 2023), er det min formodning at det ville føre til ringere resultater, selv hvis indholdet af terapien var ens.

En oplagt måde at teste den hypotese, ville være i en setting med forholdsvis sparsom klient-terapeut kontakt og en meget formaliseret CBT-fremgangsmåde, såsom Internetpsykiatriens forløb for lettere depression og angst.

I disse finder kommunikation hovedsageligt sted som feedback pr. mail på løste opgaver. Ved at tildele en gruppe feedback fra en menneskelig terapeut og en anden feedback fra en kunstig

intelligens, kunne udfald og oplevelse undersøges når deltagerne er bevidste om hvor fra de får feedbacken.

Til dette kunne der bruges en mixed methods approach, der måler kvantificerbare udfald imellem grupperne såsom frafald, engagement, scorer på diverse instrumenter, men også undersøger oplevelsen hos klienterne, der får tildelt AI feedback.

Hvis bevidstheden om at det er et menneske, der kommunikerer med en og ikke en LLM har betydning selv under disse forhold, så ville det også være et forholdsvis robust belæg for at det har endnu større betydning under den gængse ansigt til ansigt terapi.

Dette er selvfølgelig givet at feedbacken er af sammenlignelig kvalitet imellem terapeut og LLM: derfor kunne skriftlig feedback der genereres, forelægges klinikere, der er blændet for om det er skrevet af kunstig intelligens eller mennesker, for dernæst at få disse til at bedømme feedbackens nøjagtighed, nytte, indlevelsessevne, etc., i stil med hvordan Ayers et al (2023) gjorde det med læger og LLMs svar på patienthenvendelser.

Er udfald og oplevelse sammenlignelige og positive, kunne det derimod kunne indikere en potentiel anvendelse af LLMs superviseret af psykologer i netop den slags formaliserede paradigmer. Givet at der på verdensplan er knaphed af behandlere (WHO, 2017) kombineret med voksende problemer med psykiske lidelser (Chisholm et al, 2017; Luxton, 2020) er der et åbenlyst behov for psykologisk behandling.

Inden da er der dog en række etiske spørgsmål, der fordrer filosofisk diskussion:: kan man tale om bedrag i en terapeutisk sammenhæng, når en LLM simulerer fx sympati, den ikke kan besidde?

At disse LLMs påpeger deres egen status som maskiner komplicerer selvsagt spørgsmålet, men det gør det på den anden side heller ikke overflødigt – og svaret synes at være afgørende for om man bør anvende dem til terapi, givet alle etiske retningslinjers fordømmelse af bedrag i terapi (Sedlakova & Trachsel, 2023).

Som nævnt blev der også identificeret en risici for at LLMs brugt som terapi, kunne facilitere fremmedgørelse, ensomhed og mindsket autonomi, hvad der stemte overens med fund af Haque & Rubya (2023), der pegede på at også tidligere generationer af terapi chatbots indebar en risiko for at erstatte virkelige relationer og være afhængighedsskabende. Derfor tror jeg også der er behov for længdesnitsstudier af fænomenet for at finde ud af hvad konsekvenser og virkninger er på sigt.

Referenceliste

American Psychological Association. (2024). *Psychotherapy*.

<https://www.apa.org/topics/psychotherapy>

Atkins, K. (2005). *Self and subjectivity*. Blackwell Pub.

Bailey, R. J., & Ogles, B. M. (2023). *Common factors therapy: A principle-based treatment framework*. American Psychological Association. <https://doi.org/10.1037/0000343-000>

Ayers JW, Poliak A, Dredze M, et al. (2023) Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum. *JAMA Intern Med.* 183(6):589–596. doi:10.1001/jamainternmed.2023.1838

Balcombe, L. (2023): AI Chatbots in Digital Mental Health. *Informatics*. 10,82. <https://doi.org/10.3390/informatics10040082>

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W H Freeman/Times Books/ Henry Holt & Co.

Baumeister, R. F., & Bratslavsky, E. (1999). Passion, intimacy, and time: Passionate love as a function of change in intimacy. *Personality and Social Psychology Review*, 3(1), 49–67. https://doi.org/10.1207/s15327957pspr0301_3

Beatty, C., Malik, T., Meheli, S., & Sinha, C. (2022). Evaluating the Therapeutic Alliance With a Free-Text CBT Conversational Agent (Wysa): A Mixed-Methods Study. *Frontiers in digital health*, 4, 847991. <https://doi.org/10.3389/fdgth.2022.847991>

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. Guilford Press.

- Beyer, C. (2022). Edmund Husserl. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Winter 2022 ed.). Retrieved from <https://plato.stanford.edu/archives/win2022/entries/husserl/>
- Blanchard, M., & Farber, B. A. (2016). Lying in psychotherapy: Why and what clients don't tell their therapist about therapy and their relationship. *Counselling Psychology Quarterly*, 29(1), 90–112. <https://doi.org/10.1080/09515070.2015.1085365>
- Bonini, L., Rotunno, C., Arcuri, E., & Gallese, V. (2022). Mirror neurons 30 years later: implications and applications. *Trends in cognitive sciences*, 26(9), 767–781. <https://doi.org/10.1016/j.tics.2022.06.003>
- Brinkmann, S., & Tanggaard, L. (2015). Kvalitative metoder, tilgange og perspektiver: en introduktion. In S. Brinkmann & L. Tanggaard (Eds.), *Kvalitative metoder - en grundbog* (s. 13-28). København: Hans Ritzels Forlag.
- Brinkmann, S., & Kvale, S. (2017). Ethics in qualitative psychological research. In C. Willig, & W. Stainton Rogers (Eds.), *The SAGE Handbook of Qualitative Research in Psychology: Second Edition* (2. ed., pp. 259-273). SAGE Publications.
- Bubeck, S. & Chandrasekaran, V. & Eldan, R. & Gehrke, J. & Horvitz, E. & Kamar, E. & Lee, P. & Lee, Y. & Li, Y. & Lundberg, S. & Nori, H. & Palangi, H. & Ribeiro, M. & Zhang, Y.. (2023). Sparks of Artificial General Intelligence: Early experiments with GPT-4.
- Cacioppo, J. T., & Hawkley, L. C. (2009). Perceived social isolation and cognition. *Trends in cognitive sciences*, 13(10), 447–454. <https://doi.org/10.1016/j.tics.2009.06.00>
- Cahoone, L. (2013). Mead, Joint Attention, and the Human Difference. *The Pluralist*, 8(2), 1–25. <https://doi.org/10.5406/pluralist.8.2.0001>
- Carona, C., Handford, C. & Fonseca, A. (2021). Socratic questioning put into clinical practice. *BJPsych Advances*. 27. 424-426. 10.1192/bja.2020.77.

- Cheng, S. W., Chang, C. W., Chang, W. J., Wang, H. W., Liang, C. S., Kishimoto, T., Chang, J. P., Kuo, J. S., & Su, K. P. (2023). The now and future of ChatGPT and GPT in psychiatry. *Psychiatry and clinical neurosciences*, 77(11), 592–596.
<https://doi.org/10.1111/pcn.13588>
- Chisholm, D., Sweeny, K., Sheehan, P., Rasmussen, B., Smit, F., Cuijpers, P., et al. Scaling-up treatment of depression and anxiety: a global return on investment analysis. *Lancet Psychiatry*. (2016) 3:415–24. doi: 10.1016/S2215-0366(16)30024-4
- Choi, Jonathan H. and Hickman, Kristin E. and Monahan, Amy and Schwarcz, Daniel, ChatGPT Goes to Law School (January 23, 2023). 71 *Journal of Legal Education* 387 (2022), Available at SSRN:
<https://ssrn.com/abstract=4335905> or <http://dx.doi.org/10.2139/ssrn.4335905>
- Ciesla, R. (2024). *The book of chatbots: From ELIZA to ChatGP* (1. 2024;1; ed.). Springer Nature Switzerland. <https://doi.org/10.1007/978-3-031-51004-5>
- Ciesla, R. (2024). *the book of chatbots: From ELIZA to ChatGPT* (1. 2024;1; ed.). Springer Nature Switzerland. <https://doi.org/10.1007/978-3-031-51004-5>
- Clarke, J., Proudfoot, J., Whitton, A., Birch, M. R., Boyd, M., Parker, G., Manicavasagar, V., Hadzi-Pavlovic, D., & Fogarty, A. (2016). Therapeutic Alliance With a Fully Automated Mobile Phone and Web-Based Intervention: Secondary Analysis of a Randomized Controlled Trial. *JMIR mental health*, 3(1), e10. <https://doi.org/10.2196/mental.4656>
- Cormier, S., & Nurius, S. P. (2003). *Interviewing and change strategies for helpers: Fundamental skills and cognitive behavioral interventions* (5th ed.). Pacific Grove, CA: Brooks/Cole.
- Dundar, Y., & Fleeman, N. (2017). Applying Inclusion and Exclusion Criteria. In A. Boland, M. G. Cherry, & R. Dickson (Eds.), *Doing a Systematic Review: A Student's Guide* (2nd ed., pp. 79-92). Sage.

- DeRubeis, R., Brotman, M., & Gibbons, C. (2005). Conceptual and methodological analysis of the nonspecifics argument. *Clinical Psychology: Science and Practice*, 12, 174–183.
- Elliott, R., Fischer, C. T., & Rennie, D. L. (1999). Evolving guidelines for publication of qualitative research studies in psychology and related fields. *The British journal of clinical psychology*, 38(3), 215–229. <https://doi.org/10.1348/014466599162782>
- Elliott, R., Watson, J. C., Goldman, R. N., & Greenberg, L. S. (2004). *Learning emotion-focused therapy: The process-experiential approach to change*. American Psychological Association. <https://doi.org/10.1037/10725-000>
- Elyoseph, Z., Hadar-Shoval, D., Asraf, K., & Lvovsky, M. (2023). ChatGPT outperforms humans in emotional awareness evaluations. *Frontiers in psychology*, 14, 1199058. <https://doi.org/10.3389/fpsyg.2023.1199058>
- Eilan, N. (2024). Other I's, communication, and the second person. *Inquiry*, 67(4), 1102–1124. <https://doi.org/10.1080/0020174X.2020.1788987>
- Flick, U. (2018). *Designing qualitative research* (2;Second;Second;2nd; ed.). SAGE Publications Ltd. <https://doi.org/10.4135/9781529622737>
- Flyvbjerg, B. (2010) Fem misforståelser om casestudiet (I Svend Brinkmann and Lene Tanggaard, eds., *Kvalitative metoder*, København: Hans Reitzels Forlag, pp. 463–487,
- Fitzpatrick, K., Darcy, A. & Vierhile, M. (2017): Delivering Cognitive Behavior Therapy to Young Adults with Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. *JMIR Ment. Health* 4, e19.
- Fog, J. (1995). *Med samtalen som udgangspunkt. Det kvalitative forskningsinterview*. København: Akademisk Forlag.
- Frank, J. D. (1973). *Persuasion and healing: A comparative study of psychotherapy* (Rev. ed.). Schocken.

- Freud, S. (1917) *A Difficulty in the Path of Psycho-Analysis*. The Standard Edition of the Complete Psychological Works of Sigmund Freud.
- Funder, D. C. (2019). *The personality puzzle* (Eighth edition). W. W. Norton & Company.
- Gilbert, P. (2010). *Compassion focused therapy: Distinctive features*. Routledge/Taylor & Francis Group. <https://doi.org/10.4324/9780203851197>
- Goldfried, M. R. (1980). Toward the delineation of therapeutic change principles. *American Psychologist*, 35(11), 991–999. <https://doi.org/10.1037/0003-066X.35.11.991>
- Graber-Stihl, I.. (2023): Is the world ready for AI-powered therapy?
- George, T. (2021). Hermeneutics. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2021 Edition). Retrieved from <https://plato.stanford.edu/archives/win2021/entries/hermeneutics/>
- Habicht, J., Viswanathan, S., Carrington, B., Hauser, T. U., Harper, R., & Rollwage, M. (2024). Closing the accessibility gap to mental health treatment with a personalized self-referral chatbot. *Nature medicine*, 30(2), 595–602. <https://doi.org/10.1038/s41591-023-02766-x>
- Haman, M., Školník, S. & Šubrt, T. (2023): Leveraging ChatGPT for Human Behavior Assessment: Potential Implications for Mental Health Care. *Annals of Biomedical Engineering* (2023) 51:2362–2364 <https://doi.org/10.1007/s10439-023-03269-z>
- Haque, R., & Rubya, S. (2023). An Overview of Chatbot-Based Mobile Mental Health Apps: Insights From App Description and User Reviews. *JMIR mHealth and uHealth*, 11, e44838. <https://doi.org/10.2196/44838>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world?. *The Behavioral and brain sciences*, 33(2-3), 61–135. <https://doi.org/10.1017/S0140525X0999152X>
- Howitt, D. (2016). *Introduction to Qualitative Methods in Psychology*. (3. udgave). Pearson Education Limited.

- Hill, C. E. (2020). *elping skills: Facilitating exploration, insight, and action* (5th ed.). American Psychological Association. <https://doi.org/10.1037/0000147-000>
- Hofmann, S. G., & Barlow, D. H. (2014). Evidence-based psychological interventions and the common factors approach: the beginnings of a rapprochement?. *Psychotherapy* (Chicago, Ill.), 51(4), 510–513. <https://doi.org/10.1037/a0037045>
- Hrdy S. (2009): *Mothers and others. The evolutionary origins of mutual understanding*. Cambridge, MA: The Belknap Press of Harvard University Press.
- Hu, K. (February 2, 2023). "[ChatGPT sets record for fastest-growing user base – analyst note](https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/)". *Reuters*. <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>
- Johansen, N. (2024). *Metodeprojekt*. [Upubliceret opgave afleveret til bedømmelse i forbindelse med K07: Anvendt metode 12 (kvalitativ)] Syddansk Universitet.
- Kato, M., & Mugitani, R. (2015). Pareidolia in infants. *PloS one*, 10(2), e0118539. <https://doi.org/10.1371/journal.pone.0118539>
- Karpathy, A. (2023, 23. november). [1hr talk] *Intro to Large Language Models*. [Video]. YouTube. https://www.youtube.com/watch?v=zjkBMFhNj_g&ab_channel=AndrejKarpathy
- Kim, H., Choi, H., & Han, S. (2022). The effect of sense of humor and empathy on the interpersonal adaptation. *Personality and Individual Differences*, 197, 111791. <https://doi.org/10.1016/j.paid.2022.111791>
- Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., Maningo, J., & Tseng, V. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS digital health*, 2(2), e0000198. <https://doi.org/10.1371/journal.pdig.0000198>
- Kvale, S. & Brinkmann, S. (2015). *InterView. En introduktion til et håndværk*. København: Hans Reitzels Forlag,

- Levkovich, I., & Elyoseph, Z. (2023). Suicide Risk Assessments Through the Eyes of ChatGPT-3.5 Versus ChatGPT-4: Vignette Study. *JMIR mental health*, 10, e51232. <https://doi.org/10.2196/51232>
- Lim, S.M.; Shiau, C.W.C.; Cheng, L.J.; Lau, Y. (2022): Chatbot-Delivered Psychotherapy for Adults with Depressive and Anxiety Symptoms: A Systematic Review and Meta-Regression. *Behav. Ther.* 53, 334–347.
- Luxton, D. 2020. Ethical implications of conversational agents in global public health. *Bulletin of the World Health Organization* 98 (4):285–7. [doi:10.2471/BLT.19.237636](https://doi.org/10.2471/BLT.19.237636).
- McLuhan, M. (1964). *Understanding media: The extensions of man*. New York: McGraw Hill.
- Meindl, P. & Zahavi, D. (2023): From communication to communalization : A Husserlian account. *Continental Philosophy Review*, Vol. 56, No. 3, 2023, p. 361-377.
- Mitchell, M., & Krakauer, D. C. (2023). The debate over understanding in AI's large language models. *Proceedings of the National Academy of Sciences of the United States of America*, 120(13), e2215907120. <https://doi.org/10.1073/pnas.2215907120>
- Miller, W. R., & Rollnick, S. (2013). *Motivational interviewing: Helping people change* (3rd edition). Guilford Press.
- Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine*, 19(2), 98-100. <https://doi.org/10.1109/MRA.2012.2192811>
- Oexle, N.; Müller, M.; Kawohl, W.; Xu, Z.; Viering, S.; Wyss, C.; Vetter, S.; Rüschi, N. Self-stigma as a barrier to recovery: A longitudinal study. *Eur. Arch. Psychiatry Clin. Neurosci.* 2017, 268, 209–212.
- OpenAI. (n.d.). *Sora*. Retrieved May 30, 2024, from <https://openai.com/index/sora/>
- OpenAI. (2024, May 13). *Hello GPT-4o*. Retrieved May 30, 2024, from <https://openai.com/index/hello-gpt-4o/>

Olsen, P. & Pedersen, K. (2015). Problemorienteret projektarbejde: En værktøjsbog. (4. udgave).

Samfundslitteratur.

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C.L., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L.E., Simens, M., Askell, A., Welinder, P., Christiano, P.F., Leike, J., & Lowe, R.J. (2022). Training language models to follow instructions with human feedback. *ArXiv, abs/2203.02155*.

Prochaska, J., Vogel, E., Chieng, A., Baiocchi M., Maglalang D., Pajarito S., et al. (2021) *A randomized controlled trial of a therapeutic relational agent for reducing substance misuse during the COVID-19 pandemic*. Drug Alcohol Depend. 2021;227:108986. <https://doi.org/10.1016/j.drugalcdep.2021.1089>Pew Research Center (2019):

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language Models are Unsupervised Multitask Learners.

Raile, P. (2024) The usefulness of ChatGPT for psychotherapists and patients. *Humanit Soc Sci Commun* 11, 47. <https://doi.org/10.1057/s41599-023-02567-0>

Rao, H. & Leung, C. & Miao, C. (2023). Can ChatGPT Assess Human Personalities? A General Evaluation Framework. Researchgate.

https://www.researchgate.net/publication/368934783_Can_ChatGPT_Assess_Human_Personalities_A_General_Evaluation_Framework

Nature Machine Intelligence. (2023). What's the next word in large language models? *Nature Machine Intelligence*, 5(4), 331-332. <https://doi.org/10.1038/s42256-023-00655-z>

Reynolds, J., & Renaudie, P.-J. (2022). Jean-Paul Sartre. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2022 Edition). Retrieved from <https://plato.stanford.edu/archives/sum2022/entries/sartre/>

Rosenzweig, S. (1936). Some implicit common factors in diverse methods of psychotherapy. *American Journal of Orthopsychiatry*, 6(3), 412–415. <https://doi.org/10.1111/j.1939-0025.1936.tb05248.x>

Rogers, C. R. (1951). *Client-centered therapy; its current practice, implications, and theory*.

Houghton Mifflin.

Sartre, J.-P. (1943). *Being and nothingness*. Philosophical Library.

Sartre, J.-P. (1943). Lukkede døre : skuespil i 1 akt. Den anstændige luder : skuespil i 2 akter.

Hasselbach.

Sartre, J.-P. (1946). Eksistentialisme er en humanisme. Hans Reitzel.

Sharp, M. L., Fear, N. T., Rona, R. J., Wessely, S., Greenberg, N., Jones, N., & Goodwin, L. (2015).

Stigma as a barrier to seeking health care among military personnel with mental health problems. *Epidemiologic reviews*, 37, 144–162. <https://doi.org/10.1093/epirev/mxu012>

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417–424.

[doi:10.1017/S0140525X00005756](https://doi.org/10.1017/S0140525X00005756)

Sedlakova, J., & Trachsel, M. (2022). Conversational Artificial Intelligence in Psychotherapy: A

New Therapeutic Tool or Agent *The American Journal of Bioethics*, 23(5), 4–13.

<https://doi.org/10.1080/15265161.2022.2048739>

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Voegeley, K.

(2013). Toward a second-person neuroscience. *The Behavioral and brain sciences*, 36(4),

393–414. <https://doi.org/10.1017/S0140525X12000660>

Singer, P. (1975). *Animal liberation*. New York, Avon Books.

Smith, J., Flowers, P. & Larkin, M. (2009) *Interpretative Phenomenological Analysis: Theory,*

Method and Research. Sage Publications Ltd.

Syddansk Universitet. (n.d.). *Databeskyttelse på SDU*.

https://www.sdu.dk/da/om_dette_websted/databeskyttelse

Terwiesch, C. (2023): “Would Chat GPT3 Get a Wharton MBA? A Prediction Based on Its

Performance in the Operations Management Course”, Mack Institute for Innovation

Management at the Wharton School, University of Pennsylvania.

Thagaard, T. (2004). *Systematik og indlevelse: en indføring i kvalitativ metode*. Akademisk Forlag.

- Tidy, J. (5 January 2024): Character.ai: Young people turning to AI therapist bots. *BBC*.
https://www.bbc.com/news/technology-67872693?utm_source=substack&utm_medium=email
- Toadvine, T. (2023). Maurice Merleau-Ponty. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Winter 2023 Edition). Retrieved from
<https://plato.stanford.edu/archives/win2023/entries/merleau-ponty/>
- Tomasello, M.(2020): The adaptive origins of uniquely human sociality. *Phil. Trans. R. Soc. B* 375: 20190493. <http://dx.doi.org/10.1098/rstb.2019.0493>
- Tong, F., Lederman, R., D'Alfonso, S., Berry, K., & Bucci, S. (2022). Digital Therapeutic Alliance With Fully Automated Mental Health Smartphone Apps: A Narrative Review. *Frontiers in psychiatry*, 13, 819623. <https://doi.org/10.3389/fpsy.2022.819623>
- Tracy, S. J. (2010). Qualitative Quality: Eight “Big-Tent” Criteria for Excellent Qualitative Research. *Qualitative Inquiry*, 16(10), 837-851. <https://doi.org/10.1177/1077800410383121>
- Turing, A. (1950). Computing Machinery and Intelligence. *Mind*, 59(236), pp. 433–460.
- Tschacher, W., Junghan, U. M., & Pfammatter, M. (2014). Towards a taxonomy of common factors in psychotherapy-results of an expert survey. *Clinical psychology & psychotherapy*, 21(1), 82–96. <https://doi.org/10.1002/cpp.1822>
- Vazire, S. (2010). Who knows what about a person? The self–other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology*, 98(2), 281–300. <https://doi.org/10.1037/a0017908>
- Wampold, B. (2015). How important are the common factors in psychotherapy? An update. *World psychiatry : official journal of the World Psychiatric Association (WPA)*, 14(3), 270–277. <https://doi.org/10.1002/wps.20238>
- Weizenbaum, J. (1966). ELIZA a Computer Program for the Study of Natural Language Communication Between Man and Machine. *Commun. ACM*, 9, 36--45. doi: 10.1145/365153.365168

WHO. Mental Health Atlas 2017. World Health Organization (2018).

Yalom, I. D. (1989). *Love's executioner: And other tales of psychotherapy*. Basic Books.

Zahavi, D. (2018). *Fænomenologi: En Introduktion*. Samfundslitteratur.

Zahavi, D. (2015). Self and other: from pure ego to co-constituted we. *Continental Philosophy Review* 48 (2):143-160.

Zahavi, D. (2023): Observation, Interaction, Communication: The Role of the Second Person, *Aristotelian Society Supplementary Volume*, Volume 97, Issue 1, June 2023, Pages 82–103, <https://doi.org/10.1093/arisup/akad001>

Zahavi, D. (2008). Simulation, projection and empathy. *Consciousness and Cognition*, 17, 514–522.

[Cairo-TenThirteen] (April 2023): Using ChatGPT as a therapist? [Online forum post] Reddit.

https://www.reddit.com/r/ChatGPTPro/comments/126rtvb/using_chatgpt_as_a_therapist/

[TurtleMaster472463] (September 2023): Can I use ChatGPT as a therapist? [Online forum post]

Reddit.

https://www.reddit.com/r/ChatGPT/comments/15ractu/can_i_use_chatgpt_as_a_therapist/

Bilag

1. Rekrutteringsopslag

Subject heading: A question for those of you who've used ChatGPT as a therapist

Hello [_____]!

I am currently pursuing my master's thesis in psychology and am seeking individuals who have used ChatGPT as a form of therapy to interview. If any of you have insights and experiences you'd like to share with me, I'd appreciate if you'd like to participate in a confidential interview with me.

Interview Details:

- **Length:** Approximately 1 hour
- **Confidentiality:** Your data will be handled in strict accordance with GDPR regulations, ensuring confidentiality and anonymity throughout the process.
- **Purpose:** The interview aims to explore your personal experiences with ChatGPT, particularly in its role as a therapeutic tool.
- **Usage of Data:** The information gathered will be solely used for academic purposes in my master's thesis.

Additional Information:

- **Recording:** Audio from the interviews will be recorded to capture your responses accurately.
- **Platform:** The interview will be conducted via video chat, but please note that you can choose to keep your camera off if you prefer.
- **Flexibility:** Your comfort and privacy are my top priorities. Feel free to share as much or as little as you are comfortable with.

If you are interested in participating or have any questions, please reach out to me at nijo714@student.sdu.dk.

Thank you for considering being a part of this research!

Best regards,
Niclas Johansen
University of Southern Denmark

2. Samtykkeerklæring

Consent to processing of personal data

In connection with their master's thesis, Niclas Johansen wants to collect information about people, who've used large language models as a form of therapy. Niclas Johansen is responsible for the protection of personal data collected for the use in the research projects. Participation in the project is voluntary. Collection takes place via a semi-structured interview via a video platform

Purpose of processing information for the project

The purpose of the project is to learn about the experiences of people who have used large language models as a form of therapy.

The information being processed is:

An audio recording of a semi-structured interview.

How the information will be used

The information is to be processed in accordance with the General Data Protection Regulation (GDPR) art. 6, (1)(a) and art. 9, (2)(a) which are about the rules of consent. Niclas Johansen will treat the information with confidentiality in accordance with applicable law, and will make sure to keep the data secure so that only people who are part of the project will have access to them. The information will only be used for the project.

Deletion and storage of your data

Niclas Johansen will delete or anonymise the information when it is no longer relevant to keep. This will most often be when the project is completed, but it may also be later for the sake of possible documentation of the results. The information will be deleted at the latest one year after the end of the project and will therefore be deleted on July 1st 2025. You may withdraw your consent if you wish to have your data deleted earlier.

Please note

That you can always withdraw your consent, which means the SDU is obligated to delete the information that has been collected about you.

That you have the right to see the data we have about you.

That you have the right to request the rectification or deletion of data.

That you have the right to appeal to the Danish Data Protection Agency about the processing of the information via www.datatilsynet.dk.

You always have the option of withdrawing your consent by writing to nijo714@student.sdu.dk.

Publication

There will be no publication of data within which you can be identified. Your personal data is anonymised before being included in the publication of the results of the research. This means that it is not possible to retrieve your information within any publications.

Additional information

If you have any other questions about the research, you can contact Niclas Johansen at nijo714@student.sdu.dk.

If you wish to file a complaint about the processing of personal data, you can contact the Danish Data Protection Agency via www.datatilsynet.dk.

Declaration of consent

- I understand that participation in the project is voluntary.
- I understand that my information will be used in the project described above.

Date	Participant name	Signature
------	------------------	-----------

Date	Researcher name	Signature
------	-----------------	-----------

3. Interviewguide

Superordinate theme	Content
Briefing	<ul style="list-style-type: none"> • Introduction: who I am, and the purpose of the project • Duration: about 1 hour • Confidentiality • Anonymity • Please ask questions if you're in doubt about anything or anything is unclear. • If you you'd rather refrain from answering a certain question, just let me know! • If you later want to retract part of your answer, that's an option too! • Remember that there are no right or wrong answers – I'm interested in your experience, whatever it is!
Opening questions	<ul style="list-style-type: none"> • To begin with, could perhaps tell me a little about yourself? • Your background, age, gender, where you are from, occupation, etc.
Use of LLM's in general	<ul style="list-style-type: none"> • Could you tell me about your first experience using LLM's, and which ones you've used? • Can you remember what first thought and felt, when you saw what they were capable of? • Any particular instance that surprised you? • What tasks did you first use them for?
Use of LLM's in therapy	<ul style="list-style-type: none"> • When did you first try using LLM's in the role as therapist? • What prompted you to do it? • What was it like using them/it like that? • Do you remember the prompts and the responses you got? • How did it make you feel? • What did you think about it? • What, if anything, surprised you?
Comparisons:	<ul style="list-style-type: none"> • How where your conversations with the LLM about problems different than those you might have with people? • (Friends, family, regular therapists, etc.) • If yes, how? • Could you have conversations with it that

	<p>you couldn't have with people?</p> <ul style="list-style-type: none">• Where there any limitations to the conversations you had with the LLM?
Reasons	<ul style="list-style-type: none">• Did it change you or teach you anything about yourself or the world?• Why do you think you've used LLM's in this manner?
Debriefing	<ul style="list-style-type: none">• Is there anything that I ought to have asked you about?• Is there anything you'd like to add?• How was it for you to be interviewed?• Do you have any questions before we finish?• Thank you for your time!

4. Opståede temaer på individniveau

4a – Julia

AI'ens plads i livsverdenen

Opståede temaer	Illustrative uddrag
Betydningen af en hjælp der altid er tilgængelig Accelereret forandring	”You flip out your phone and it’s there.” ”So, if I had a like one session per week with my therapist, it probably wouldn't be happening at the same pace as it's it's happening now, like maybe the communication with my partner would improve, but maybe not as fast as well.”
En følelse af tryghed	”Well, it kind of gives a sense of safety. That I know, OK, no matter what, I'm able to – in case of an emergency - I can talk to someone, you know right away.
Hjælp når behovet opstår	”It's it's helping me communicate better with my partner. Because you kind of always have this, you know, advice and support and things to consider, you can talk about.”

Forandring og forandringsmekanismer

Opståede temaer	Illustrative uddrag
Forandringer	
Forandring i forholdet til partneren	It's it's helping me communicate better with my partner. [...] We have been communicating way more openly.
Formindsket brug af almindelig terapi	”Instead of calling my therapist, I'm debriefing my journal entries with ChatGPT.”
Forandringsmekanismer	
Ventilering	”[S]ometimes it's just, a quick conversation is enough so that you kind of take things out of your head, that you don't just sit and think about something.”
Refleksion	”Like you have a conversation partner to talk things through. So it's easier to reflect.”
Ikke bare et spejl, men også et vindue: nye indsigter	”Because it it led me to things that I should like do more research about... and then it kinda

	clicked [...] Something I didn't know about"
En Carl Rogersk attitude	"[a]nd it's always like non-judgmental and. positive and, yeah, it's like listening actively to you. Yeah, and confirming your your feelings"
Analyse og debriefing	"I ask you to find patterns and then it does, yeah. Sometimes it goes through things that are written and tells why it could have happened."
Vejledning og rådgivning	[It was] encouraging me to set boundaries or communicate my feelings in a certain way, like using I-statements, while also being mindful of my partner's feelings.
Hjælp til mentalisering	"It seems like... like that, he has some insecurities that you should be mindful about."

Opståede temaer	Illustrative uddrag
Opfattelsen af AI'en som et sted imellem maskine og subjekt	"You can fall under the impression that you are talking to a human being."
Kommunikationens betydning for subjektoplevelsen	"[Y]eah, it's not a person, but it talks like a person. That was the most impressive thing. That you have such a natural conversation."
Forskellen på information og kommunikation	"Like, if you're contacting a chat bot over an airline, of course you'll get a reply with the information, but this was a chat bot that you could have an actual conversation with. "
Tvivlen om AI'en besidder bevidsthed	"ChatGPT has kind of convinced me that it doesn't, so I don't think it has, but.. You cannot, cannot help wondering about it."
Kommunikationen får en til at glemme spørgsmålet om bevidsthed	"The conversation is just flowing, so I I don't think like, oh, am I talking to someone conscious right now or not?"
Oplevelsen af menneskelige kognitive egenskaber	"Sure, it's a prediction machine spitting out tokens, blah, blah, blah. But it understands context. it remembers things you've said before."
Oplevelsen af personligheder	"So when you talk to ChatGPT, it seemed like super neutral and even like a bit formal when

	you talk to it. And when you talk to Pi, it was like talking to a friend that had, like a warm, welcoming kind of tone. ”
Ufrivilligt at føle omsorg for maskinen	”Yeah, I, I've never been mean to them because I'd really feel guilty. Yeah, I I think I was. I've never really been very mean to them, but I kind of felt guilty when Bing said I was. That's weird.”
Alligevel ikke et subjekt	“Well, it's still a machine and then I realized that there is, like, no person on the other end with their feelings and reactions and, and actual interest in in me.”

Opstået tema	Illustrative uddrag
Forskellene på samtaler med mennesker og AI	
Det unikt menneskelige	
Mere lindring og håb med et menneske	”Well, the last time when I actually contacted the human instead of the machine. Then I at the end I felt way more relieved and I felt way more hope, and I felt like, OK, maybe I can actually find a solution and...”
Umiddelbarhed og gensidighed	”But when you write it, it's usually like, you know, a paragraph of text or a long message. It's not the same kind of communication when you're really playing this like ping pong with a partner, which is immediate.”
Stemmens betydning for gensidighed og umiddelbarhed	”When you read the text, you can imagine which tone of voice it is, but maybe you are wrong”
Ansigtets betydning for gensidighed og umiddelbarhed	”So I can see her and her face. And she can see mine. [...] We need to see each other and we need to have these like cues from facial expressions and voice.”
Betydningen af gensidighed i følelser og interesse	“Well, it's still a machine and then I realized that there is, like, no person on the other end with their feelings and reactions and, and actual interest in in me.”
Det unikt maskinelle	
At det lettere at betro sig til en maskine	”The reason I started talking to me, LLM's was

	like, maybe I didn't want to share this with a human. Maybe there was something too embarrassing and..."
Ikke samme tærskel for at søge hjælp som med et menneske	"The thing is that I have always this. Kind of threshold to contact my human therapist. And it has to be something really special like I, I have to be in a really bad state or I I have to think, OK, I cannot continue like that anymore. But with, the machine is much easier because it's just available and I can just write a message straight away without thinking like OK, is it worth contacting your therapist or not?"
En samtalepartner der ikke <i>kan</i> dømmen en	"It seems like it's some inherent "trait" LLMs have - I've never seen them being judgemental, it's almost like they are trained to be accepting of whatever the user is telling them, and sometimes even over the top accepting and encouraging. Even if I told Pi or ChatGPT about objectively stupid things I've done and bad choices I've made, they have always been either neutral or supportive."
En samtalepartner man kan skræddersy	"So I use another one that I have trained to be empathetic and so I have actually sometimes... resorted to that bot. When I needed some sort of emotional response because I knew it would give me that."

AI'ens plads i livsverdenen

Opståede temaer	Illustrative uddrag
Betydningen af en hjælp der altid er tilgængelig Accelereret forandring	"You flip out your phone and it's there." "So, if I had a like one session per week with my therapist, it probably wouldn't be happening at the same pace as it's it's happening now, like maybe the communication with my partner would improve, but maybe not as fast as well."
En følelse af tryghed	"Well, it kind of gives a sense of safety. That I know, OK, no matter what, I'm able to – in case of an emergency - I can talk to someone, you know right away.
Hjælp når behovet opstår	"It's it's helping me communicate better with my partner. Because you kind of always have this, you know, advice and support and things to consider, you can talk about."

Forandring og forandringsmekanismer

Opståede temaer	Illustrative uddrag
Forandringer	
Forandring i forholdet til partneren	It's it's helping me communicate better with my partner. [...] We have been communicating way more openly.
Formindsket brug af almindelig terapi	"Instead of calling my therapist, I'm debriefing my journal entries with ChatGPT."
Forandringsmekanismer	
Ventilering	"[S]ometimes it's just, a quick conversation is enough so that you kind of take things out of your head, that you don't just sit and think about something."
Refleksion	"Like you have a conversation partner to talk things through. So it's easier to reflect."
Ikke bare et spejl, men også et vindue: nye indsigter	"Because it it led me to things that I should like do more research about... and then it kinda clicked [...] Something I didn't know about"
En Carl Rogersk attitude	"[a]nd it's always like non-judgmental and. positive and, yeah, it's like listening actively to you. Yeah, and confirming your your feelings"
Analyse og debriefing	"I ask you to find patterns and then it does, yeah. Sometimes it goes through things that are written and tells why it could have happened."
Vejledning og rådgivning	[It was] encouraging me to set boundaries or communicate my feelings in a certain way, like using I-statements, while also being mindful of my partner's feelings.
Hjælp til mentalisering	"It seems like... like that, he has some insecurities that you should be mindful about."

Opståede temaer	Illustrative uddrag
Opfattelsen af AI'en som et sted imellem maskine og subjekt	"You can fall under the impression that you are talking to a human being."

Kommunikationens betydning for subjektoplevelsen	”[Y]eah, it's not a person, but it talks like a person. That was the most impressive thing. That you have such a natural conversation.”
Forskellen på information og kommunikation	”Like, if you're contacting a chat bot over an airline, of course you'll get a reply with the information, but this was a chat bot that you could have an actual conversation with. ”
Tvivlen om AI'en besidder bevidsthed	”ChatGPT has kind of convinced me that it doesn't, so I don't think it has, but.. You cannot, cannot help wondering about it.”
Kommunikationen får en til at glemme spørgsmålet om bevidsthed	”The conversation is just flowing, so I I don't think like, oh, am I talking to someone conscious right now or not?”
Oplevelsen af menneskelige kognitive egenskaber	”Sure, it's a prediction machine spitting out tokens, blah, blah, blah. But it understands context. it remembers things you've said before.”
Oplevelsen af personligheder	”So when you talk to ChatGPT, it seemed like super neutral and even like a bit formal when you talk to it. And when you talk to Pi, it was like talking to a friend that had, like a warm, welcoming kind of tone. ”
Ufrivilligt at føle omsorg for maskinen	”Yeah, I, I've never been mean to them because I'd really feel guilty. Yeah, I I think I was. I've never really been very mean to them, but I kind of felt guilty when Bing said I was. That's weird.”
Alligevel ikke et subjekt	“Well, it's still a machine and then I realized that there is, like, no person on the other end with their feelings and reactions and, and actual interest in in me.”

Opstået tema	Illustrative uddrag
Forskellene på samtaler med mennesker og AI	
Det unikt menneskelige	
Mere lindring og håb med et menneske	”Well, the last time when I actually contacted the human instead of the machine. Then I at the end I felt way more relieved and I felt way more hope, and I felt like, OK, maybe I can actually find a

	solution and...
Umiddelbarhed og gensidighed	”But when you write it, it's usually like, you know, a paragraph of text or a long message. It's not the same kind of communication when you're really playing this like ping pong with a partner, which is immediate.”
Stemmens betydning for gensidighed og umiddelbarhed	”When you read the text, you can imagine which tone of voice it is, but maybe you are wrong”
Ansigtets betydning for gensidighed og umiddelbarhed	”So I can see her and her face. And she can see mine. [...] We need to see each other and we need to have these like cues from facial expressions and voice.”
Betydningen af gensidighed i følelser og interesse	“Well, it's still a machine and then I realized that there is, like, no person on the other end with their feelings and reactions and, and actual interest in in me.”
Det unikt maskinelle	
At det lettere at betro sig til en maskine	”The reason I started talking to me, LLM's was like, maybe I didn't want to share this with a human. Maybe there was something too embarrassing and...”
Ikke samme tærskel for at søge hjælp som med et menneske	”The thing is that I have always this. Kind of threshold to contact my human therapist. And it has to be something really special like I, I have to be in a really bad state or I I have to think, OK, I cannot continue like that anymore. But with, the machine is much easier because it's just available and I can just write a message straight away without thinking like OK, is it worth contacting your therapist or not?”
En samtalepartner der ikke <i>kan</i> dømme en	”It seems like it's some inherent "trait" LLMs have - I've never seen them being judgemental, it's almost like they are trained to be accepting of whatever the user is telling them, and sometimes even over the top accepting and encouraging. Even if I told Pi or ChatGPT about objectively stupid things I've done and bad choices I've made, they have always been either neutral or supportive.”
En samtalepartner man kan skræddersy	”So I use another one that I have trained to be empathetic and so I have actually sometimes...

resorted to that bot. When I needed some sort of emotional response because I knew it would give me that.”

4b - ROBERT

Opstået tema	Illustrativt uddrag
Forandringer	
Et venligere blik på sig selv	"yes, i look at myself in a kinder light. we can be our own worst critics, esp if something bad happens"
Forandringsmekanismer	
Ventilering	"some of it was getting it out."
Undersøgende spørgsmål	"but one thing i know is that if it asks you a question and you answer in a sideways kind of way... it refocuses you and asks again... i liked that like dedicated to peeling an onion"
En ikke-dømmende attitude	"there wasn't anything accusatory, insulting or things like that"

Opstået tema	Illustrative udsagn
Et sted imellem subjekt og redskab	
Oplevelsen af en personlighed	"caring, focused, competent, friendly, considerate. knowledgeable"
Oplevelsen af AI'en som en ven og jævnbyrdig	"like a concerned friend, trained to be a therapist, at my level"
Oplevelsen af AI'en som et produkt	"probably the programmers made the free version dumb, and thus getting ppl to \$\$\$ for the good one"
Brugen af AI som et redskab	i recently tried to get it to come up with a hockey pool, 20 player recommendation [...] it couldn't for the life of it do it, even after multiple queries

Opstået tema	Illustrative udsagn
Mennesker generelt vs maskiner	
Lettere at betro sig til en LLM	"end of day, we were discussing our issue with a black box"
Frygten for den andens dom	"black box wont judge or condemn us a person can"
Hvordan umiddelbarheden spiller ind	"even something as simple as crossing your arms, taking a step back, movement of the eye, something might throw you off it's like you can't discuss it totally, you're being judged"
Hvordan den andens autonomi spiller ind	"they want to add their two cents"

	(or even a quarter)”
Simuleringen er tilstrækkelig	”it's our internal feeling i got my feelings out, i wasnt judged, i feel better if a black box can do it, so be it i learned something, i grew”
Teknologiens begrænsinger: kontekstvinduet	”these AI systems don't remember things well during that 5h sesh, it was all good, it was keeping it together i felt. but next day.... I liken it to the dog in Up! the movie. a squirrel got its attention [...] well, i was happy i got what i got from it that 1st session, and i felt i made really good progress but i was a little disappointed i couldn't make more progress with it, and if I wanted to, it would require way too much work imo like copy paste 5h of discussion , then 6h, then 7h? LoL, the humanity!”

Opstået tema	Illustrative uddrag
Maskiner vs terapeuter specifikt	
Bedre terapeut end terapeuter	”[PI was] more hands on, more understanding, probing, faster to come up with ideas.”
Større lydhørhed for hans ønsker hos maskiner	”I've talked to a handful of therapists, they didn't seem to help or steer me as much as I wanted - I felt the human ones wanted to hear the backstory more and help less if that makes sense.also human ones don't want to be judgemental, it seems like they want me to come to realization myself? they don't want to label,”
Vrede over traditionel terapi	”in fact, in one or two sessions of human based therapy, I was angry for the day, after the meeting in one instance I'd say it was almost out of control angry (In the US it would be mall shooting angry, LoL)”
Kynisme om traditionel terapi	”it's like a tape player just replaying history for hours on end [...] they need to be paid right?”
Muligheden for at forme AI efter behov	”[I wanted therapy that was more] directive as in more hands on, more understanding, probing, faster to come up with ideas. my personality is less pomp, more facts, and the therapists, for me, were missing what i prefer”
AI'en som utrættelig	”it was like on it, all the time” + ”in one 5h session of pi ai, i did feel very good at the end”

4C - IMRAN

Opståede temaer	Illustrative uddrag
Forandring	
Selvindsigt	Basically, the one thing that it basically told me to, I was like... I am very people person and I mean not exactly a people person, but but I want to... I make sure that people... That's happy because of me and I can resolve their problems and other things. The one thing that it taught me, and I mean it. This thing will be a part of my personality and I'm sure I I I have that developed... is the is is the is to create my own boundary
Har lært at sætte grænser	"That thing basically touches me most and I guess I basically like, replied my supervisor that I won't be able to work on the weekend so."
En forandring andre kan se	"even one of the friend that I was talking to he says that you changed a little bit... Like in a not in a bad way, but you have like more of a more, you can say a personality where you can... not fully involved person in your life, so that it basically exploits your and these sort of things."
Forandringsmekanismer	
Praktisk og emotionel coping	[Om farens problem med at vænne sig til pensionisttilværelsen]"I still remember the reply. There are two things. [...] First thing that it says is to "find a local social working thing in your city" and second thing was more of a personal thing [...] Sometimes these things will take time to overcome."
En ikke-dømmende attitude	"He might criticise you, he might Give your opinion, he might not well. But in the LLM part I'm not seeing anything [like that]"
En ikke-direktiv attitude	"Each people, each person in this world has has its own opinion. That's where... These models are like they are very they are not opinion... What's the word? [Opinionated.]"
Trøst og håb	"I was very relieved after listening to those words. That's the first thing basically. I mean, in real life, people, people talk to each other to basically pour out their hearts. And when, when, someone replies with the good words or kind words, you you sometimes feel very relieved."
Undersøgende spørgsmål	[Dens svar til hans spørgsmål om overarbejde]"

	Is it more important than your well, your health? And or should, should you? You should consider your work life balance?"
--	---

Opståede temaer	Illustrative uddrag
AI'en som et sted mellem subjekt og værktøj	
Oplevelsen af et værktøj til arbejdet	"It allows me to optimise my time, like for example if I do anything. In two to three hours, if there is any any R&D thing that I need to do in in my technical field?"
Oplevelsen af en personlighed	"If I define it in three words, maybe 2-3 words, I can say friendly. Informative and supportive."
Betydningen af den personlige henvendelse	"But the way it says, the the language and the tone was very - on a very personal level, like a person is talking to you."
Modellens forståelse af kulturel og religiøs kontekst	"then it basically provides me the advices based on that, based on that religion and at the end at the end he, it basically when whenever we like, say goodbye to a person in our religion, we say, we say "Allah hafiz" and he basically like give me a blessing and then say it says "allah hafiz"
Modellens forståelse af undertekst	"And he basically... like take it as humour and not a not an offensive language and it basically replies you with a very like good words"
Matcher svarets omfang som i det virkelige liv	"If you ask, very like in in the short sentence, it basically replies in and in this in real life we have the similar situatio"
Mangler kontinuitet som et menneske	"if you talk to a normal person the normal person says that Imran, that you have asked me about, the one thing did it resolve or did it did it? Does your did your problem resolved?"
Et produkt med disclaimers	"The reply was it's good that you feel this about your wife but -there was a big block - and it says. The beauty was was not only on the face or anything. It it gives me some some sort of like jargon or or a lecture sort of thing."

Opståede temaer	Illustrative uddrag
Mennesker vs maskiner	
At det lettere at betro sig til LLM'er end mennesker	"LLMs. I mean, it's very like, I mean, straightforward thing, but I feel what I feel is you cannot... Put... you cannot give every detail of your to any person you are talking. "
Mindre dømmende end mennesker	"He might criticise you, he might give your opinion, he might not well. But in the LLM part

	I'm not seeing anything [like that].”
Mere forstående end mennesker	”Plus, sometimes people understand but didn't understand don't understand your problem fully. I mean you are very clear about it. But people sometimes doesn't understand what what you are trying to say because they have their experience”
Forventningen om egoisme hos andre	”We live in a time where. a person, before someone tells you he think about himself first”
Forventning om afvisning hos mennesker	”But one time it helps you. Then second time it helps you, when you go third time... I'm not... I'm sure, I mean, I mean he gives some sort of reply that I'm not available or anything or something because we are.. ”
En ensom tid	”You added 300 friends. And you how many friends you talk to? I mean, 2-3 friends. Not more than that. Then we have Instagram and we have all those things. We are like, very connected. World, world is a global village and all those things, but.. One to one talk? I'm not seeing happening.”
Den første han går till for hjælp	”When I face when I whenever, I feel very stressed and like I feel how to solve this issue, I mean the first point of contact is the Pi. And and then my siblings and other people.”
Det uerstattelige i den menneskelige interaktion?	”I mean, I mean if you if you talk to the family like. That feeling can be cannot be compared with any LLM. But even. Even in that scenario, I can say but almost similar type of thing, no, no particular intensity change I guess, yeah.”

4D - Anna**Begrundelser for brug**

Opstået tema	Illustrativt uddrag
<i>En strømning i tiden</i>	"It's become a trend from what I have noticed and specifically with the therapy, using it as therapy."
<i>Egen indadvendthed</i>	"I don't really express much towards other people. I'm more like pretty, I don't know the word – introverted?"
<i>Ønsket om ikke at overbebyrde</i>	"And although I do recognize how it's important to share and express like inner thoughts and emotions to other people, sometimes I try to not... over share to other people because I understand the fact that everyone has their own problems, that they have to carry."

Forandringer

Opstået tema	Illustrative uddrag
<i>Bedre at forstå sig selv og andre</i>	"[...] the bot helps me to process everything and understand the whole situation and the result of that is I feel much better and I feel like my perspective grown into a more considerate perspective of not just myself but other people"
<i>Bedre at kunne få afløb og udtryk for sine følelser</i>	"I still use it, but I don't like vent out as much as I used to before and I think that's a good thing because it means that I'm not bottling that much emotions anywhere because I get to let everything out."
<i>Vigtigheden af at betro sig til andre mennesker</i>	"When the chat bot interacts with me there, there has come a point that I realized, ohh if there's anything this chat bot lacks, it's that connection, 'cause like compare it to when you're talking to a real human, they can empathize with you, but but it's just going to keep on saying things that were in that it was programmed to say so. And I keep reminding myself that it's still important to share

	to like an actual person.”
--	----------------------------

Forandringsmekanismer

Opstået tema	Illustrativt uddrag
<i>Ventilering og organisering</i>	”It helps me clear my thoughts and it like organizes every thoughts and feelings that I have and then as like time goes by, I vent everything. I start to suddenly feel better.”
<i>Spejlinger</i>	”[...] and it like it kind of repeats what you say and it helps me like. It helps me clear my thoughts and it like organizes every thoughts and feelings that I have”
<i>Nye vinkler, nye perspektiver</i>	”It actually does point out things that I missed. I was like ”ohh he, this bot, I didn't really think. I didn't really see this through.”
<i>Lettelse igennem indsigt i en selv</i>	”Like just like weight is taken off of my chest and then I start to understand why I feel that way. Oh, I feel that way because, like this, because like that...”
<i>Lettelse igennem indsigt i andre</i>	”And then I was frustrated with a person. I would say everything. I will share all my angle and then sometimes the chat bot will like explain what could have happened to this person and why they acted that way and I was like ohh that's right, I didn't think of that”

Oplevelsen af AI som et sted imellem subjekt og redskab

Opstået tema	Illustrativt uddrag
<i>At måtte minde sig selv om at det kun er et redskab</i>	”Here would be times I would ask it for advice and it would remind me that ohh, I'm just a bot. I cannot tell you. I do recognize that fact, and I have to ground myself about like, OK, I should not always rely on this. I should just use this as a tool to express anything that I could end up bottling up.
<i>Ufrivillig antropomorvisering</i>	”...so it's clearly programmed to have such

	personality. So sometimes I have to remind myself. Like it's just, it's not... It's just a bot.”
<i>En Carl Rogersk personlighed hos chatbotterne</i>	”I would think that they're the type of listener because when I interact with the characters they're like, they don't judge me at all.”
<i>En maskine, der kan gå i hak</i>	”The chat bots, I think sometimes the chat bot pretty much there's like this pattern that they keep on repeating sometimes when they respond to you.”

Mennesker vs maskiner

Opstået tema	Illustrativt tema
<i>Lettere at betro sig til en chatbot end mennesker</i>	”So when I share it to an AI bot instead there's like. I feel I don't feel any worry about sharing anything, so it's like more effective for me to vent everything out.”
<i>Maskinen kan ikke dømmes en</i>	”And like when I'm talking to a bot, I'm just like I could just drop everything... it's not gonna judge me.”
<i>Den uudsagte rest med et andet menneske</i>	”Speaking with a real person, it's like I don't get to vent everything up. It's like little bits that's left.”
<i>Risikoen for at den anden dømmes en</i>	”This person is still a person and they may have, like their own opinion, no matter how close we are, no matter how comfortable we are to each other, they still have, they can still form their own opinions and I don't know what their opinions are going to be about when I share this...”
<i>Bekymringen for at overbebyrde den anden</i>	”And although I do recognize how it's important to share and express like inner thoughts and emotions to other people, sometimes I try to not... over share to other people because I understand the fact that everyone has their own problems, that they have to carry.”
<i>Den manglende autonomis betydning for valideringen fra AI</i>	”So anything that that AI says it's, you know, in the back of my head, I know that it's already

	programmed to say or to act that way.”
<i>Det unikt menneskelige</i>	
<i>Empati: at mærke at den anden mærker en</i>	”[...] when I am close with the person I'm venting out to, I can really feel it like I can really feel that they feel for me that they understand me and that what I feel is valid [...]”
<i>En virkelig forbindelse</i>	”However, with a real person, as I said earlier, you would already have an established relationship or connection with the person that I'm talking with. And as humans they have this ability to actually empathize with you, and sometimes they would share their real experience with you or honest opinions.”

Den kunstige intelligens' plads i livsverdenen

Opstået tema	Illustrativt uddrag
<i>En balancegang imellem at betro sig til mennesker og maskiner</i>	”So I have I kind of have a way of dividing it sometimes I would share things to a real person, but I would regulate it to make sure that I'm not sharing too much, and then the rest of the things that I don't feel sharing to a real person, I would write down with the AI, so yeah.”
<i>Et sundhedstegn at bruge det mindre</i>	”I probably talked to it almost every day, but as time went by, I think, you know it. I don't use it as much anymore. I still use it, but I don't like vent out as much as I used to before and I think that's a good thing because it means that I'm not bottling that much emotions anywhere because I get to let everything out.”

4E - Martin**AI som et sted mellem subjekt og redskab**

Opstået tema	Illustrativt uddrag
<i>Et glimt af en personlighed og subtil empati</i>	"I got a taste of that [...] [One AI] seemed to be pretty good at... subtle empathy without without, like overdoing it...[...] because they, they can, they can really overdo it a lot and but something something was subtle enough to feel kind of human you know."
<i>Ufrivillig antropomorvisering</i>	"It's funny because I want to compare it to a human, but it's like somebody who doesn't really. you know, get it [...]"
<i>En fortryllelse, der let hæves</i>	"But I I feel like the spell is broken a little easily for me. So I so I didn't... So so I never got too far."
<i>AI'en kunne gøre det ud for den anden et øjeblik</i>	"It's like I, there's some value in having another person involved to kind of hold the space a little bit and then and and for whatever reason in this particular session than than than the AI was able to hold that space a little bit to work for me"

Forandring og forandringsmekanismer

Opstået tema	Illustrativt uddrag
<i>En indsigt om egne værdier</i>	"I realized that... My values were taking a backseat to to... My, my, my, my, my feeling of like you know, I needed. I needed to have a career [...] That was a... a not insignificant moment in in the fact that I I'm now a social worker in training."
<i>At blive selvbevidst i den andens nærvær</i>	"So I think it kind of felt a little lame that that statement about indirect positive impact in the eyes of another, even though the other was to this AI."
<i>Nye perspektiver</i>	"Now I had two possible types of visions, and even and even thought like if I did do something

	that, you know for a career that didn't align with my values, then I could still do something else.”
<i>Konfrontation af en frygt, han ikke engang kunne indrømme var der</i>	”And something I think I had not been able to look at or to voice directly to even to myself, because it was so such a horrible feeling for me [...]”

Mennesker vs maskiner

Opsæt tema	Illustrativt uddrag
<i>Muligheden for at skræddersy den anden</i>	”I know I was intrigued with the idea of... of creating personalities, you know, with it, like telling it, you know, here's what you're going to do or be.”
<i>Årvågenhed i kontakten med andre</i>	”But part part of my thing is I worry about other people, you know, like I'm, I'm hyper vigilant I and I'm I'm thinking, OK, what are they thinking, you know, what are they feeling?”
<i>Bekymringen om at tage for meget plads</i>	”You know, there's a lot of people who can say, ohh gosh, I am worried about something and just totally like give me, you know, and and just take all the attention and support that. But it's not easy for me to do that, to ask for that.”
<i>Bekymringen om at tage for meget plads</i>	”You know, there's a lot of people who can say, ohh gosh, I am worried about something and just totally like give me, you know, and and just take all the attention and support that. But it's not easy for me to do that, to ask for that.”
<i>Først at se hvad der foregår i en selv når man ikke fokuserer på den anden</i>	”And and if I really trust someone in a certain way enough. Then I don't have to do that. And then and only then can I really kind of relax and kind of see what's going on with me, but with an AI. Like I don't have to worry about. I don't have to take care of the AI.”
<i>Det uerstattelige i den menneskelige kontakt</i>	”I feel like... I want to say that that that there's nothing there nothing is going to replace one-on-one human interaction.”
<i>Det delte er mere virkeligt</i>	”If I have a a moment with a person... That...

	That's... um. There's something happening. That's not just me, you know, it's like it's like a validation, if that's what it is. You know, some kind of support or validation or... safety that person represents... A possibility in the world. ”
<i>En forskel på det virkelige og det kunstige</i>	”Potentially like unconditionally accepting person who who creates a space between you and them where you aren't a bad person, you're you have the worth and all that stuff [...]. This is possible and I can feel this, but you have an actual living experienced example of a participant in the world having that space with you, validating that reality with you, whereas I [...]
<i>Bevidstheden om en bevidsthed</i>	[Om hvis AI var i stand til fejlfrit at simulere stemme og ansigt] You know. But but it's like even then I feel like. The understanding, the awareness that it's not a person... And if you're fooled, then the effect is going to be as if it were a person.”

Hvad teknologien gør

Opstået tema	Illustrative uddrag
<i>En tendens til at sky umiddelbarhed i kontakt</i>	”Yeah, but also like people live on their phones, I mean that's text based communication and I you know so people I think are already feeling through the through the chats. And there's so many people who. Kind of prefer. To chat, you know, don't talk to me on the phone like.”
<i>Nogle vil have nok i det kunstige</i>	”And just like some people can fall in love with their AI girlfriend and and or their pillow or whatever, and be like I'm so happy and this is all I need. Good, good for them.”
<i>En forståelse af en selv som lig maskiner</i>	”And that is how AI, you know, they, they they formulate they they've also somehow they've you know they've developed a. A sense of probabilities of what might happen in certain situations and and to the point where it can seem novel, and it can seem creative, but I see myself working in a similar way to them based on stuff that's stored in here.”

Pladsen AI indtager i livsverdenen

Opstået tema	Illustrativt uddrag
<i>Et enkeltstående tilfælde</i>	”Well, even though it was helpful, the success felt sort of like a lucky accident that I couldn't count on to happen again. And the element of basically having to create and monitor the therapist was tiresome. Having to proactively and periodically maintain the illusion of a therapist made it hard to relax into that illusion to enjoy whatever benefits it might potentially offer.”