



**LUNDS**  
UNIVERSITET

Institutionen för psykologi  
*Psykoterapeutprogrammet*

**Exponeringsbaserad VR-behandling för fobier:  
en systematisk litteraturöversikt av  
forskningskvaliteten**

**Henrik Altréus & Tony Kaveh**

Psykoeterapeutexamensuppsats. 2021

Handledare: Philip Lindner

Examinator: Sean Perrin

## Abstract

Over the years, there have been several Randomised Controlled Trials (RCTs) supporting the effectiveness of Virtual Reality Exposure Therapy (VRET) in the treatment of phobias. However, there has not been enough scrutiny of the overall quality of these publications. In this study, 7 VRET RCTs between 2013 and 2019 were reviewed using Cochrane as well as Öst's criteria for research design. No one had previously reviewed these studies using both criteria. Conclusions from the study is that surprisingly few quality RCT publications are available so far in this emerging field. Based on the criteria used in our study however, the publications were of high quality measured by both scales. The need for other reasonable and more reality-based research review criteria is discussed.

*Keywords:* virtual reality, VRET, exposure therapy, phobia, anxiety, social.

## Sammanfattning

Under årens lopp har det gjorts flera randomiserade kontrollerade studier (RCT) som stödjer effekt av VRET (Virtual Reality Exposure Therapy) vid behandling av fobier. Det har verkar dock ha saknats tillräcklig granskning av den övergripande kvaliteten på dessa studier. I den här studien undersöktes forskningskvalitet hos sju VRET RCTs publicerade mellan 2013 och 2019 med hjälp av Cochrane samt Östs verktyg för forskningsdesign. Ingen hade tidigare granskat dessa studier med utifrån bägge uppsättning kriterier. Den övergripande skattningen påvisade hög kvalitet utifrån båda utvärderingsverktygen. Dessutom blev vi förvånade över att vi inte hittade fler RCT-publicationer av hög kvalitet inom detta nya område än vad vi hade antagit. Vår slutsats är att detta kan bero på VR:s snabba utveckling och differentiering av forskningsfälten. Behovet av andra rimliga och mer verklighetsbaserade kriterier för granskning av forskning diskuteras.

*Keywords:* virtual reality, VRET, exposure therapy, phobia, social, anxiety

Exponeringsbaserad VR-behandling för fobier: .....	1
en systematisk litteraturstudie av forskningskvalitet .....	1
Fobier.....	1
Epidemiologi och prevalens .....	2
Behandling av ångestsjukdomar och fobier .....	2
<i>Kognitiv beteendeterapi</i> .....	2
<i>Virtual reality baserad behandling</i> .....	2
<i>Tidigare forskning</i> .....	3
Metod .....	4
Resultat.....	5
Resultat av sökning och selektion. ....	5
Deskriptiv och kvalitativ sammanfattning av studier .....	7
Bedömning med Cochrane risk of bias tool .....	10
<i>Selection bias</i> .....	12
<i>Performance bias</i> .....	12
<i>Detection bias</i> .....	12
<i>Attrition bias</i> .....	13
<i>Reporting bias</i> .....	13
Kvalitativ bedömning med Psychotherapy rating scale .....	13
<i>Studiedesign, allokering och jämförbarhet</i> .....	16
<i>Utfallsmått</i> .....	16
<i>Statistisk bearbetning och signifikansbedömning</i> .....	16
<i>Beskrivning, kronicitet och representativitet hos sample</i> .....	16
<i>Bedömning och reliabilitet</i> .....	17
<i>Samtidiga behandlingar</i> .....	17
<i>Mätpunkter för utfallsmått</i> .....	18
<i>Evidensbaserad interventioner</i> .....	18
<i>Terapeuters antal och kompetens</i> .....	18

<i>Kontroll av följsamhet till behandlingen</i> .....	18
<i>Bortfall och statistiska analyser</i> .....	19
<i>Kvantativ översikt av medelvärden</i> .....	19
<i>Helskalans poäng per artikel</i> .....	20
Diskussion .....	22
Summerande diskussion och kritisk reflektion .....	24
Referenser .....	25

## **Exponeringsbaserad VR-behandling för fobier: en systematisk litteraturstudie av forskningskvalitet**

Även om farmakologiska behandlingar för psykisk ohälsa har vunnit mer mark och fått mer uppmärksamhet genom åren, är de inte fullt ut accepterade av alla patienter, och en del kommer att uppleva biverkningar eller inte svara tillräckligt bra på dem (Harvery, 2015; Fournier, 2010). Vidare har vi evidensbaserade psykologiska behandlingar /interventioner som patienter kan erbjudas (som enda eller kompletterande behandling), som har visat lovande resultat och gör det möjligt för patienterna att lära sig färdigheter som hjälper dem att hantera sina symtom samtidigt som de kan förebygga eventuella risker för återfall (Dellazizzo, 2020). Dock har effekten av psykoterapier för psykiska störningar i bästa fall visat sig vara måttlig och med ett bortfall och avbrottsfrekvens på ca 30 %. Behandlingsvinsterna av terapin lär därför inte alla gånger bibehållas på lång sikt (Harvery 2015; Swift 2014).

Med tanke på de ovanstående bristerna i den konventionella behandlingen finns det alltså fortfarande ett behov av nya terapiformer som ska adressera dessa tillkortakommanden. Således har dessa försök till att öka effektiviteten, acceptansen och tillgången till evidensbaserade psykoterapier lett till framväxten av en del teknikbaserade psykologiska interventioner. Ett utmärkt exempel på detta är *Virtual reality* (VR) som vi kommer att beskriva och diskutera i det här arbetet.

### **Fobier**

Fobiska ångeststörningar (ICD-10 F40) klassificeras som en underkategori av ångeststörningar i ICD-10 och kännetecknas av ångest i avgränsade situationer som för närvarande utgör en liten eller ingen faktisk fara som leder till ett undvikande av dessa situationer eller det som man är rädd för (Världshälsoorganisationen, 1992). Det finns tre subtyper av fobiska ångeststörningar i ICD-10; agorafobi (F40.0), social fobi (även kallad SAD) (F40.1) och specifik fobi (F40.2). Patienter med specifik fobi är rädda för specifika situationer eller föremål som djur, höjder, mörker eller slutna och trånga utrymmen. Patienter med social fobi brukar vara rädda för att bli granskade av andra människor, vilket leder till att de undviker sociala situationer. De med agorafobi (torgskräck) är en fobi är rädda för att besöka offentliga platser eller hamna i situationer där det inte är så lätt fly från situationen, till exempel i olika sammanhang där det finns folksamlingar i offentliga utrymmen, gå in i affärer eller att resa på egenhand osv. Personer med tandvårdsfobi undviker vanligtvis vanliga tandbehandlingsförfaranden och utrustning som till exempel extraktioner, tandborrar och nålar (Oosterink, F. M., de Jongh, A., & Aartman, I. H., 2008). Undvikande av tandvård har en

negativ inverkan på den orala och allmänna hälsan genom att minska tandläkarbesöken och utnyttjandet av tandvård (a.a.).

### **Epidemiologi och prevalens**

Wechsler (2019) hänvisar till flertal studier (Bijl et al., 1998; Alonso et al., 2004; Gustavsson et al., 2012; Kessler et al., 2005; Stein et al., 2017; Wardenaar et al., 2017; Stagnaro et al., 2018) och rapporterar att tolv månadersprevalensen för fobiska ångeststörningar ligger mellan 0,3 - 1,6 % för agorafobi, 1,2 - 6,8 % för social fobi och 3,5 - 8,7 % för specifik fobi. Livstidsprevalensen har rapporterats ligga mellan 0,9 - 3,4 % för agorafobi, 2,4 - 7,8 % för social fobi och 7,7 - 10,1 % för specifik fobi enligt samma källa/or (Wechsler, 2019). Enligt Global Burden of Disease Study (GBD, 2017) rankas ångestsjukdomar som den nionde största bidragande orsaken till global funktionsnedsättning, vilket ledde till totalt 24,6 miljoner som levde med funktionsnedsättning på grund av det i världen år 2015 (Vos et al., 2016).

### **Behandling av ångestsjukdomar och fobier**

#### ***Kognitiv beteendeterapi***

Kognitiv beteendeterapi (KBT) anses vara en av de mest effektiva behandlingarna av ångeststörningar (Craske et al., 2014). KBT har många olika komponenter varav en är *exponeringsterapi* (ET). Det finns en hel del empiriska bevis som talar för exponeringsterapis effektivitet, även som enda behandlingsmetod vid behandling av ångestsjukdomar och fobier (Craske et al., 2014; Davidson et al., 2004). ET genomförs genom in vivo och imaginära exponering, vilket innebär konfrontation med fruktade (men objektivt sett säkra stimuli) situationer eller minnen. De mest erkända mekanismerna bakom exponeringsterapis effektivitet är habituering, korrigerande av negativa övertygelser och emotionell bearbetning (Foa och Kozak, 1986; Clark, 1999).

***Virtual reality baserad behandling.*** På senare år har det dock blivit allt vanligare att använda sig av virtuella stimuli för att genomföra exponeringsterapi, med s.k. virtuell verklighet eller *virtual reality* (VR). Användningen av multisensorisk VR som är den mest vedertagna termen, har föreslagits som ett kostnadseffektivt och logistiskt bekvämt kliniskt verktyg för ET, i t ex förhållande till traditionella in vivo-exponerings förfaranden (Rothbaum et al., 2006). I litteraturen kallas detta oftast för *Virtual Reality Exposure Therapy* eller helt enkelt VRET. Detta har också föreslagits som en exponeringsteknik för de som har svårt att på ett effektivt sätt aktivera rädslonätverk vilket anses nödvändigt för att uppnå en terapeutisk effekt (Difede & Hoffman, 2002). VR innehåller datorgrafik, visuella skärmar och sensoriska inflöden för att skapa en engagerande och uppslukande *virtuell* miljö som befrämjar upplevelse

av närvaro i den virtuella världen. Med tanke på att VR gör det möjligt att skapa skräddarsydda virtuella miljöer lämpar sig denna teknik och modalitet väl för även för ET.

### ***Tidigare forskning***

Tre metaanalyser som genomförts om VRET för ångeststörningar har kommit fram till att VRET är överlägsen väntelistkontroll och ingen skillnad i förhållande till aktiva behandlingar (Oprish et al., 2012; Parsons & Rizzo, 2008; Powers & Emmelkamp, 2008). På samma sätt drog Meyerbroeker och Emmelkamp (2010) slutsatsen i en översikt att VRET är en lovande behandling för en stor bredd av ångeststörningar; dock med kritik gällande litteraturbasen och kanske starkare metoder. Ytterligare problem har tagits upp om kvaliteten på den nuvarande VRET-forskningen på grund av användningen av små urvalsstorlekar och en brist på bredd och enhetlighet i rapporteringen av data (Parsons & Rizzo, 2008).

Även om det har uttryckts farhågor om kvaliteten på VRET-studier, har kvaliteten på studier inom det (litteraturen) ännu ha bedömts på ett systematiskt sätt tillräckligt mycket. I en tidigare studie granskades litteraturen om psykoterapi för depression (Cuijpers, van Straten, Bohlmeijer, Hollon & Andersson, 2010) och fann att de studier som bedömdes vara av hög kvalitet var följande rapporterade mindre behandlingseffektstorlekar jämfört med studier av låg kvalitet. Författarna (a.a.) drog slutsatsen att även om effekterna av psykoterapi för depression fortfarande är signifikanta, metaanalyser har överskattat effekterna av denna intervention. Cuijpers et al. (2010) menade att denna överskattning till stor del beror på de "otillräckligt rigorösa metoder" som förekommer i litteraturen. I den här litteraturöversikten ämnar vi att utforska hur forskningskvalitén för sentida studier inom exponeringsbaserad VR-behandling ser ut och belysa detta med hjälp av sedvanliga kriterier inom etablerad psykoterapiforskning.

Då VR-behandling ännu inte blivit tillgänglig för patienter i särskilt omfattande utsträckning och återkommande kritik framförts gällande just metodologiska frågeställningar, ämnar vi utvärdera hur just forskningskvalitén ser ut och utvecklas inom exponeringsbaserad VR-behandling (VRET) för fobier. Vi väljer liksom i tidigare forskningsansatser att granska själva omfattningen av tillgänglig publikation av god kvalitet med fokus på kvalitativ och deskriptiv relevans för psykoterapiforskningen, snarare än effektstorlekar. Vi vill på så sätt förstå hur studier som gjorts på området ser ut och utvecklats över tid och utifrån dessa utgångspunkter belysa hur denna typ av forskning står sig jämfört med kriterier vanliga inom etablerad psykoterapiforskning. Förväntan är att omfattning av högkvalitativ VRET-forskning både har ökad i omfång och tillgänglighet och också utvecklats utifrån kritik som framförts



inom fältet. Vi hade också som utgångspunkt att forskningskvalitet borde visa sig relativt god då vi selekterar utifrån höga artikelkrav redan vid inklusion.

### **Metod**

Vi genomförde en systematisk litteraturoversikt med syfte att finna och granska de senaste och mest relevanta, RCT:er (Randomized Controlled Trials) gällande VRET för specifika fobier respektive social fobi. Vi valde att inte selektera mellan olika fobier utan fokusera på att undersöka och belysa den övergripande, metodmässiga forsknings kvalitén inom fältet. Som framkommer i selektionskriterier nedan utgick vi från krav om randomiserade studier med aktiv kontrollgrupp. En etikdeklaration enligt rådande regler vid Lunds universitet fylldes i av författarna och godkändes av både handledare och examinator. Sökningar genomfördes av författarna, var för sig, och synkroniserades i slutgiltig sammanställning.

Databaser som användes vid den systematiska sökningen var; PsycINFO, PubMed, och Web Of Science. Sökorden bestod av följande ordkombinationer: ”Virtual OR Phobia” och ”Virtual OR Anxiety OR social” i antingen titel, abstract eller innehåll (beroende av databasens sökalternativ). Målet var att fånga upp alla relevanta VR studier gällande både specifika fobier och social ångest. Sökningen begränsades till artiklar som var på engelska och svenska.

Eftersom Institutionens ramar kring litteraturstudier begränsade antalet artiklar till 7–10 st. valde vi vid slutlig selektering att prioritera de senaste och mest aktuella publicerade studierna, oavsett storlek, så länge de föll under kategorin för diagnosområdet, dvs specifik fobi och socialfobi. Följande inklusionskriterier användes vid urval: 1) minst en VRET-baserad behandlingskomponent, 2) randomiserad fördelning av försökspersoner och 3) aktiv kontrollgrupp. Följande kriterier gällde vid exklusion av studier: studier med ”within study”, cross-over design; studier som kombinerade VRET med annan behandlingskomponent eller intervention (psykoterapeutisk eller farmakologisk) som gjorde att VRET inte kunde utvärderas separat; studier där behandlingsdos med VRET uppenbarligen var lägre/divergerade jämfört med kontrollgrupp; och studier där det inte fanns tillräcklig statistiska data att tillgå. De artiklar som gick vidare till slutgiltig kvalitetsgranskning genomarbetades först individuellt av författarna med hjälp av Cochrane Risk of Bias Tool (Cochrane Methods, 2021) samt Psychotherapy outcome study methodology rating scale (Tolin, 1999) (svensk och generellt anpassad översättning av Lars-Göran Öst, se bilaga 1).

Inledningsvis genomförde vi en kvalitativ sammanställning och beskrivning av de för psykoterapiforskning relevanta egenskaperna i de inkluderade studierna som ett underlag för diskussion om kvalitetsbedömning av området.

För att belysa metodmässighet och risk för bias i de enskilda studierna använde vi version 2 av Risk of Bias tool (RoB 2) från Cochrane Collaboration (Higgins et al., 2019). Risken för bias inom varje område bedömdes som antingen låg, okänd, eller hög bias i enlighet med denna modifierade och något förenklade variation av Risk of bias tool. För att bedöma risken för selektionsbias, performancebias, detektionsbias, attritionbias och rapportingsbias gjordes bedömning av oss individuellt och synkroniserade resultaten genom jämförande diskussion.

För att göra en kvalitetsbedömning mest relevant för uppsatsens frågeställning, använde vi en av Östs modifierade skattningsskala för psykoterapistudier (Psychotherapy outcome study methodology rating scale (Tolin, 1999). Skalans interna konsistens är beskriven som god, med en Cronbachs  $\alpha$  på 0,86. Skalan innefattar 22 delskalor för bedömning av kvalitetsmässighet inom psykoterapiforskning. Varje delskala bedömdes utifrån tre kvalitetsnivåer: 0=Dåligt, 1=Måttlig och 2=Bra. Vidare presenterades sammantagen tabell summerad av samtliga artiklar per delskala för övergripande summering samt rangordnad poängtabell.

## **Resultat**

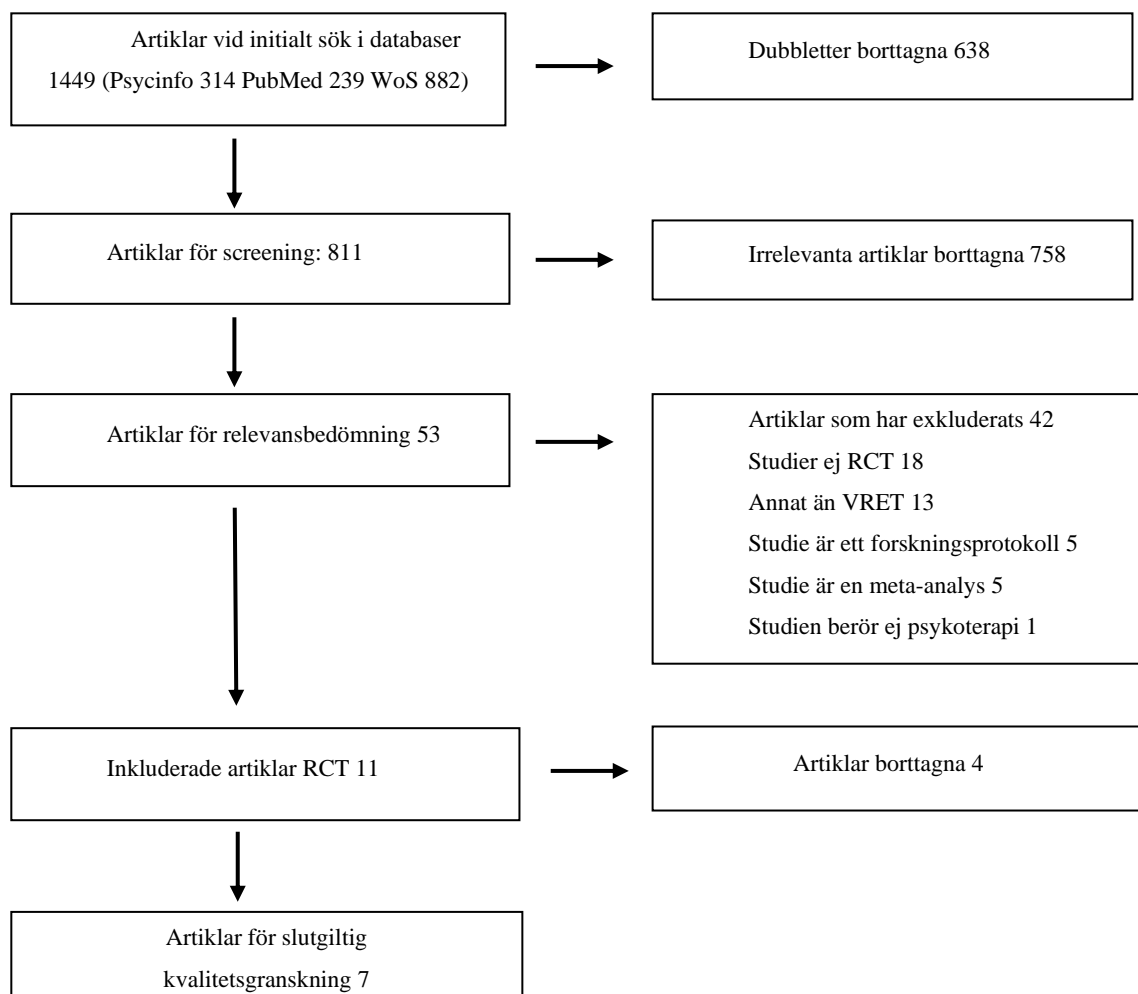
### **Resultat av sökning och selektion.**

Sökning i PubMed, PsycInfo och Web Of Science gav totalt 1449 artiklar genom söksträngar. Uppdelat på databasnivå för respektive sökning erhöles för sökning mot specifik fobi ("virtual and phobia") 715 resultat. Motsvarande för social ångest ("virtual and social and anxiety") gav 734 resultat. Resultat överfördes till EndNote där 638 dubletter togs bort. Vi gjorde även manuell kontroll för att undvika felaktig eliminering av programvara. 811 artiklar screenades sedan för övergripande tematisk forskningsdesign. 53 av dessa artiklar gick vidare till relevansbedömning och genomläsning av abstract och i vissa fall artikel i sin helhet. Av dessa uteslöts 42 studier baserat på metodmässigt uppsatta kriterier: icke RCT:s (n=18), annan typ av VR-behandling än just VRET (n=13), olika studieprotokoll (n=5), meta-analyser som inte tidigare upptäckts (n=5) samt övrig forskning icke relevant för psykoterapi (n=1). 11 RCT studier uppfyllde helt inklusions- respektive exklusionskriterier som var uppsatta.

Utifrån tidigare beskrivet grundkrav för uppsatsarbete om maximalt antal artiklar, valde vi att eliminera studier utifrån en tidsaxel så att samtliga artiklar som publicerats innan 2013 togs bort (n=4). Det fanns ett publiceringsglapp mellan studier gjorda innan 2013, vilket gjorde tidsspannet ovan rimligt att utgå ifrån, dvs från forskning gjord så mycket som möjligt i närtid. Sju studier uppfyllde kriterier för att bli inkluderade i för vidare granskning. Flödesschema för sökning och selektion framgår av Figur 1.

**Figur 1.**

*Flödesdiagram över artikelsökning och urvalsprocess*



## **Deskriptiv och kvalitativ sammanfattning av studier**

Sju studier publicerade mellan 2013 och 2019 uppfyllde kriterier för att bli inkluderade för vidare bedömning av forskningskvalitet och bias. Urvalet gav en artikel om agorafobi (Meyerbroeker et al., 2013), fyra om social fobi (Anderson et al., 2013; Bouchard et al., 2016; Kampmann et al., 2016; Lindner et al., 2019) och två om specifik fobi; dentalfobi (Gujjar et al., 2019) och spindelfobi (Miloff et al., 2019). Samtliga behandlingsupplägg föll under exponeringsinriktad behandling enligt inklusionskriterier och var samtliga individualterapeutiska. Det fanns stor variationsbredd i val av behandling med totalt sett evidensbaserade behandlingskomponenter, dock med olika teoretiska ramverk och behandlingsprotokoll. Sessionslängd pendlade också stort mellan få sessions-behandlingar till ett större antal tillfällen. Många av de redovisade primära utfallsmåtten återkom i de olika studierna men långt ifrån alla. Exponeringsbetingelser mellan experiment och kontrollgrupper var genomgående jämförbara när det kommer till miljöer, gradvis exponering, etcetera. Teknisk VR-utrustning och de program som användes varierade stort mellan grupperna, vilket snarast är att förvänta pga. stor skillnad mellan publicerade årtal, kontext för studien och all teknisk utveckling som skett under denna tid. Undantag är studier publicerade av Lindner et al (2019) respektive Miloff et al (2019) där det ser som i stort sett använder samma tekniska utrustning. Se Tabell 1 för sammanfattning av studiernas egenskaper

**Tabell 1***Kvalitativ summering av studier med aktiv kontrollgrupp inkluderade i litteraturstudien.*

Artikel	Kliniskt tillstånd	Deltagare	Primära utfallsmått	KBT-intervention	Antal sessioner	Typ av VRET	VR-utrustning
Meyerbroecker et al., 2013	Agorafobi	55	ACQ <sup>a</sup> PDSS <sup>b</sup> BSQ <sup>c</sup>	KBT-protokoll av Craske & Barlow anpassat för VR.	10	Exponering för bl.a. virtuell tunnelbana och köpcenter.	nVisor
Anderson et al., 2013	Social fobi (SAD)	97	PRCS <sup>d</sup> FNE-B <sup>e</sup> BAT <sup>f</sup>	Evidensbaserade KBT-interventioner för SAD, primärt exponering.	8	Virtuellt konferensrum och auditorium med reaktion från publik	VFX-headset
Kampmann et al., 2016	Social fobi (SAD)	60	LSAS-SR <sup>g</sup> FNE-B	Modifierade KBT-protokoll för SAD	10	Presentation och social interaktion inför varierande virtuell publik.	nVisor SX
Bouchard et al., 2017	Social fobi (SAD)	59	LSAS-SR FNE-B	Modifierat protokoll av Clark&Wells för SAD	14	Tala inför publik under möte, jobbintervju och presentation, samt hantera negativ kritik.	eMagin z800
Gujjar et al., 2019	Specifik fobi (Dental)	30	MDAS <sup>h</sup> DFS <sup>i</sup>	En-sessionbehandling med exponeringshierarki	1	Exponering för fem olika tandläkarspecifika VR-miljöer.	Oculus Headset

Lindner et al., 2019	Social fobi	50	PSAS <sup>j</sup> LSAS-SR	VR-anpassad och modifierad ensessionsbehandling samt uppföljande exponeringsträning	1	Nivåbaserat och patientstyrt VR-program för talångslan	Samsung Gear VR headset
Miloff et al., 2019	Specifik fobi (Spindel)	100	BAT	VR-anpassad one-session treatment	1	Nivåbaserat och patientstyrt VR-program.	Samsung Gear VR headset

---

Förtydligande av använda utfallsmått: <sup>a</sup>ACQ=Agoraphobic Cognitions Questionnaire, <sup>b</sup>PDSS=Panic Disorder Severity Scale, <sup>c</sup>BSQ=Bodily Sensation Questionnaire,

<sup>d</sup>PRCS= The Personal Report of Confidence as a Speaker, <sup>e</sup>FNE-B= Brief Fear of Negative Evaluation Scale, <sup>f</sup>BAT= Behavioral Approach Test, <sup>g</sup>LSAS= Liebowitz Social Anxiety Scale Self-Rated, <sup>h</sup>MDAS=Modified Dental Anxiety Scale, <sup>i</sup>DFS=DentistryFearScale och <sup>j</sup>PSAS=Public Speaking anxiety scale.

### **Bedömning med Cochrane risk of bias tool**

För att i detalj medologisk kvalitet, mer specifikt utvärdera risk för bias, använde vi verktyg från Cochrane Collaboration (Higgins et al., 2011).

Se Tabell 2 för illustration av bedömningsresultat och i text redogörelse för själva bedömningsprocessen.

**Tabell 2***Summering av Bedömning av Risk för Bias för inkluderade Studier <sup>a</sup>*

<u>Studie</u>	<u>Risk för selektionsbias</u>		<u>Risk of performance bias</u>	<u>Risk of detection bias</u>	<u>Risk of attrition bias</u>	<u>Risk of reporting bias</u>
	Random sequence generation	Allocation concealment	Blinding of participants and researchers	Blinding of outcome assessment	Incomplete outcome data	Selective reporting
Meyerbroeker et al., 2013	Oklar	Oklar	Oklar	Oklar	Låg	Låg
Anderson et al., 2013	Låg	Låg	Oklar	Oklar	Oklar	Låg
Kampmann et al., 2016	Låg	Låg	Oklar	Låg	Oklar	Låg
Bouchard et al., 2017	Låg	Låg	Oklar	Oklar	Låg	Låg
Gujjar et al., 2019	Låg	Låg	Hög	Oklar	Låg	Låg
Lindner et al., 2019	Låg	Låg	Oklar	Oklar	Låg	Låg
Miloff et al., 2019	Låg	Låg	Oklar	Oklar	Låg	Låg

<sup>a</sup> För att bedöma risk för bias i de enskilda studierna användes Risk of bias tool från Cochrane Collaboration. Varje område bedömdes som antingen Låg, Oklar, eller Hög risk för bias i enlighet med den modifierade och något förenklade versionen av (Higgins et al., 2011)



### ***Selection bias***

Gällande risken för *selektionsbias* så använde sig alla studier av slumpmässig fördelning, dock beskrevs det inte alla studier det specifika tillvägagångssättet för att generera slumpmässiga sekvenser och hur man ämnade dölja allokeringen till behandlingsbetingelser. Det finns en oklar risk för selektionsbias i en studie, eftersom metod för slumpmässig sekvens och blinda tilldelningen inte specificerades närmare. (Meyerbroeker., 2013). Bouchard et al (2017) uppgav en tilldelning baserad på slumpnummertabeller och en dold tilldelning som inte specificerades närmare än så. En annan studie rapporterade tilldelning baserad på en datoriserad slumpnummegerator och deltagande av en tredje studiekoordinator för att säkerställa en oklar tilldelning innan deltagarna rekryterades (Anderson et al., 2013). Lindner (2019) genomförde randomisering i omgångar inför inklusion genom tilldelning av slumpgenerator i block av varierande storlek. Miloff et al (2019) använde samma system, med tillägg att en andra forskare tog vid efter att process med slumpgenerator var avklarad.

### ***Performance bias***

För att *performance bias* helt skall kunna undvikas måste både deltagare och forskare vara blinda för behandlingsbetingelser, vilket inte är möjligt generellt i psykoterapiforskning och inte alls under VR- respektive in vivo-exponering. Tolerans för denna grund för bias är som vi ser det oundviklig i denna typ av studier. Eftersträvansvärt är att blinda deltagare fram till påbörjan av behandlingsbetingelse så långt detta går, vilket också gjordes i fyra studier (Kampman et al., 2013, Meyerbroeker et al., 2013, Miloff et al., 2019 och Linder et al., 2019). Gujjar et al (2019) beskriver att både deltagare och kliniker känner till allokering, vilket ger hög risk för performance bias i vår skattning, men det är oklart om de också egentligen blindat fram till allokering och får en för hård bedömning. Andra studier specificerade inte närmare tidpunkten för avblindning och därför bedöm risken för prestationsbias i samband med förbehandling oklar.

### ***Detection bias***

Vid *outcome assessment bias* när det gäller vedertagen användning av självskattningsformulär är även detta svårt att hålla blindat för deltagare vid eftermätning, då de i psykoterapiforskning oftast vet vilken grupp de tillhör. I två studier kunde dock blindning under förhandsbedömningen genomföras som en approximation, beskrivet av (Kampmann et al., 2016), och båda bedömdes som låg risk för upptäckt av bias under de beskrivna omständigheterna. Andra studier specificerade inte närmare om förskattning utfördes blint och bedömdes därför som en oklar risk för detection bias.

### ***Attrition bias***

Risk för *attrition bias* är i allmänhet låg för randomiserade studier av hög kvalitet, särskilt om dessa gör anspråk på intent-to-treat och således tar höjd för avhopp i behandlingen vid statistisk analys. De artiklar som inte redovisade detaljerad information om detta gav oklar bedömning för attrition bias (Bouchard et al., 2016 och Gujjar., 2019).

### ***Reporting bias***

Risk för *reporting bias* bedöms som *låg* i samtliga studier då använda utfallsmått var specificerade och beskrivna innan studierna påbörjades.

### **Kvalitativ bedömning med Psychotherapy rating scale**

I denna genomgång av kvalitetsgranskning med Psychotherapy rating scale har vi redovisat resultat från både skattningar i sig utifrån de individuella artiklarna, argument för dessa skattningar samt övergripande summera helhetsintryck från både hel- respektive delskala för att kunna belysa frågeställning gällande de inkluderade studiernas forskningskvalitet. Textmässig summering av kvalitetsbedömning gjorde från ett granskade resonemang kring studierna, snarare än att följa delskalor i inbördes ordning. Se Tabell 3 för illustration av kvalitetsgranskning. Den delskala som är aktuell för respektive del i beskrivningen anges med (siffra) för att förtydliga skattningsprocessen i texten.

**Tabell 3***Utvärdering av forskningskvalitet för de inkluderade artiklarna med Psychotherapy Rating Scale <sup>a</sup>*

Delskalor <sup>b</sup>	Meyerbroeker et al	Anderson et al	Kampmann et a.	Bouchard et al	Gujjar et al	Lindner et al	Miloff et al
1. Beskrivning av samplet	1	2	2	2	1	1	2
2. Kronicitet hos störningen	1	1	1	2	1	1	1
3. Representativitet	1	2	1	1	1	1	1
4. Reliabilitet för diagnosen	1	2	1	1	1	1	2
5. Specifika utfallsmått	2	2	2	2	1	2	2
6. Reliabila utfallsmått	2	2	2	2	1	2	2
7. Blindade bedömare	1	2	1	0	1	0	1
8. Träning av bedömare	1	2	1	1	0	1	1
9. Fördelning till behandling	1	1	1	1	1	2	1
10. Design	2	2	2	2	1	2	2
11. Statistisk power	0	2	2	0	1	1	2
12. Mätpunkter	1	2	1	1	1	2	2
13. Behandlingsmanual	2	2	2	1	1	2	2
14. Antal terapeuter	1	2	1	1	0	2	2
15. Terapeutträning	1	1	1	1	0	1	1
16. Följsamhet till manual	1	1	2	1	0	1	1

Delskalor <sup>b</sup>	Meyerbroeker et al	Anderson et al	Kampmann et a.	Bouchard et al	Gujjar et al	Lindner et al	Miloff et al
17. Terapeutkompetens	1	2	2	1	0	1	1
18. Samtidiga behandlingar	1	1	1	1	1	1	1
19. Handhavande av bortfall	2	2	2	1	1	2	2
20. Statistiska analyser	2	2	2	2	2	2	2
21. Klinisk signifikans	1	1	2	1	1	1	2
22. Terapitimmar	2	2	2	2	1	2	2

<sup>a</sup> För att göra en kvalitetsbedömning mest relevant för uppsatsens frågeställning, används en av Öst modifierade skattningsskala för psykoterapistudier (Psychotherapy outcome study methodology rating scale (Tolin, 1999). Varje delskala bedömdes utifrån tre kvalitetsnivåer: 0=Dåligt, 1=Hygglig och 2=Bra. <sup>b</sup> Förkortningar av delskalor har använts i tabellen och korrekt benämning av delskalorna redovisas här i ordningsföljd: (1) tydlighet i beskrivningen av sample, (2) svårighetsgrad/kronicitet, (3) urvalets representativitet, (4) tillförlitligheten hos diagnosen i fråga, (5) resultatmåttens specificitet, (6) resultatmåttens tillförlitlighet och validitet, (7) användning av blinda bedömare, (8) utbildning av bedömare, (9) fördelning av behandling, (10) utformning, (11) effektanalys, (12) bedömningspunkter, (13) manualbaserade, replikerbara och specifika behandlingsprogram, (14) antal terapeuter, (15) terapeututbildning/erfarenhet, (16) kontroll av att behandlingen följs, (17) kontroll av terapeuternas kompetens, (18) kontroll av samtidiga behandlingar, (19) hantering av bortfall, (20) statistiska analyser och presentation av resultat, (21) klinisk betydelse, (22) lika många behandlingstimmar.

### ***Studiedesign, allokering och jämförbarhet***

Delskalan studiedesign (10), som är grundläggande för en RCT, ger betyget *bra* för samtliga studier. Fördelning till behandling (9) ger även denna delskala naturligt hög skattning då randomisering till behandlingsbetingelser har gjorts per våra inklusionskriterier. Dock är det endast Miloff et al (2019) som klarat betyget *bra* som ett resultat av tydligt motiverat och angivet antal terapeuter som skyddar mot vinkling i de olika behandlingsgrupperna. Kan hända har man i övriga studier tagit höjd för detta, men det redovisas inte och drar ner betyget till *hyggligt* av den anledningen. Ett tredje område som i en RTC bör vara givet en design med randomiserat upplägg, är jämförbarhet i terapitimmars (22). Denna skala ger fem studier oavsett inriktning betyget *bra*. Man har både designmässig insikt och i artikel beskrivit väl hur man balanserar behandlingsdos mellan experiment och kontrollgrupp.

### ***Utfallsmått***

Specifika (5) och reliabla/valida (6) utfallsmått förekommer i samtliga artiklar och får betyget *bra*. De flesta av dessa utfallsmått används redan för att utvärdera psykoterapeutiskt behandlingsarbete och traditionell exponering för ångestsjukdomar och fobier, och lämpar sig väl för motsvarande VR-baserad sådan. Samtliga mått redovisas på ett tillfredställande sätt, även motivation för valet av just dessa instrument samt psykometriska egenskaper. Instrument som i vissa av studierna nyttjas för att mäta både s.k. ”cyber sickness” såväl som exempelvis upplevelse av närvaro i virtuella program, är relevanta för att utvärdera korrelation, effektstorlekar och lämplighetsaspekten i denna typ av behandling. I vår utvärdering är dock inte viktiga för diskussionen kring val av utfallsmått och forskningskvalitet.

### ***Statistisk bearbetning och signifikansbedömning***

Vidare har statistiska analyser (20) planerats väl i samtliga studier och klarar utan svårighet betyget *bra* då de innehåller både medelvärden och standardavvikelser redovisat. Presentation av genomförd statistisk analys av signifikansnivåer (21) för åtminstone primära utfallsmått finns även detta med i alla artiklar. Dock är det endast tre studier (Andersson., 2013; Kampmann., 2016 och Miloff., 2019) som klarar högsta betyg *bra*, då de redovisat användandet av Jacobsons (1991) högt satta kriterier för klinisk signifikans. Dessa kanske har använts i övriga studier, men de framkommer i så fall inte på ett tydligt sätt och ger resterande studier *hyggligt* betyg.

### ***Beskrivning, kronicitet och representativitet hos sample***

För att få högsta skattning gällande tydlig beskrivning av sample (1) för terapiforskning måste utöver inklusions- och exklusionskriterier och demografi även prevalensen för komorbida störningar redovisas. Detta gjordes i fyra studier (Anderson et al., 2013, Kampmann

et al., 2016, Bouchard et al., 2017 och Miloff et al., 2019). Övriga studier nådde upp till *hygglig* nivå på denna delskala, vilket minst är att förvänta hos RTC-studier. Kronicitet hos störningen (2) bedömdes för sex artiklar till nivå *hygglig* av en samma anledning då de inte klarar krav satt för kroniskhet över >1år. Undantaget var Bouchard et al (2017) där grundkrav hos kroniskhet hos försökspersoner i deras studie skulle överstiga 2 år. Representativitet hos försökspersoner (3) är viktig för bibehållen extern validitet och generaliserbarhet. Denna delskala är relativt vag i sin beskrivning och således mer öppen för tolkning gällande de mycket olika rekryteringsförhållanden som råder i studierna. Övergripande har alla studier utom en fått bedömning nivå *hygglig*, vilket betyder medvetenhet och strävan efter att urvalet ska vara generaliserbart. Vid läsning av Kampmann et al (2016) framkom att rekryteringen helt gjordes via annonser utan samarbete med vårdorganisationer, vilket gör det svårt att helt se försökspersoner i deras studie som representativa för patienter som söker vård för denna problematik. Anderson et al (2013) som vi gav skattning *bra*, hade i kontrast både samarbete med olika vårdgivare för sin rekrytering och tog dessutom kulturella aspekter i beaktning vid kartläggning och redovisning av demografi.

### ***Bedömning och reliabilitet***

När det kommer till reliabilitet vid bedömning av försökspersoner (4) så fann vi i beskrivning av förlopp i samtliga studier strukturerade intervjuer som var genomförda av utbildad bedömare, vilket nådde upp till grundkrav *hyggligt* betyg på denna delskala. Anderson et al (2013) såväl som Miloff et al (2019), skattade och redovisade också interbedömarreliabilitet utöver skattning av erfaren bedömare, vilket krävs för nivå *bra* som vi bedömde lämplig i deras studier. Det var antingen otydligt beskrivet i studier av Bouchard et al (2017) och Lindner et al (2019) huruvida de ämnade eller hade använt blindade bedömare (7), och detta gav omdöme *dåligt* i detta fall. Anderson et al (2013) var den enda studien som beskrev testning av inspelade sessioner på ett slumpmässigt sätt för att avgöra om bedömare var medvetna om behandlingsbetingelser, vilket krävs i skalan för att nå högsta kvalitetsbedömning och betyget *bra*. De övriga artiklarna beskrev förekomst av blindhet, men utan tester för att kontrollera detta. Gällande bedömares utbildning och specificering av dennes träning (8) fick samtliga artiklar utom Anderson et al. (2013) *hygglig* skattning då beskrivning eller förekomst av kalibrering mätprocedurer saknades.

### ***Samtidiga behandlingar***

Kontroll av samtidig behandling (18) gjordes på en *hygglig* nivå i samtliga studier. Kriterier för att uppnå bra skattning på denna skala är hel frånvaro av annan psykologisk eller medicinsk behandling. I samtliga artiklar bibehöll man krav om avslutad psykologisk

behandling, men var mer flexibla gällande stabilitet i dosering av medicin, vilket gav betyget hyggligt.

### **Statistisk poweranalys**

Statistisk poweranalys (11) för att bedöma och anpassa samplestorlek gjordes i tre studier med bra skattning (Andersson., 2013; Kampmann,; 2016 och Miloff., 2019), hyggligt i två studier (Gujjar., et al 2019) och Lindner et al., 2019) och slutligen dålig skattning i de två sista studierna (Buchard et al., 2017 och Meyerbroeker.,et al 2013).

### ***Mätpunkter för utfallsmått***

Användandet av flera mätpunkter (12) före, efter och vid en uppföljande tidpunkt är en vedertagen praktik i psykoterapiforskning och i RCT-studier. Detta gjorde också att samtliga studier minst klarade hyggligt betyg. Långtidsuppföljning tidigast efter 12 månader krävs dock för att en studie skall få betyget bra på denna delskala, vilket endast uppnåddes i tre artiklar (Anderson et al., 2013, Lindner et al., 2019, Miloff et al., 2019).

### ***Evidensbaserad interventioner***

När det kommer till evaluering av kvalitet i behandlingsrelaterade program som användes (13) var det endast en av artiklarna (Buchard et al., 2017) som inte helt levde upp till krav om tillräckligt detaljerad beskrivning, dvs replikerbarhet. I övriga studier hade man antingen utgått från existerande behandlingsprogram och generellt ändrat delar som hade med exponeringen att göra, eller använde väl beskrivna och replikerbara evidensbaserade interventioner anpassade till störningen, och gav därmed betyg bra.

### ***Terapeuters antal och kompetens***

Vid skattning av antal terapeuter som genomförde behandlingen (14) var det tre studier (Anderson et al., 2013, Lindner et al., 2019 och Miloff et al., 2019) som klarade kraven för *bra* betyg genom att både minst ha två terapeuter per behandlingsbetingelse i studien samt analysera varje terapeuts påverkan på utfallet. Övriga studier utom en fick betyg *hyggligt*, då den sistnämnda delen med analys saknades. Undantag gällande avvikande kvalitetsmässighet när det kommer till beskrivning av kliniska insatser är tydligt hos Gujjar et al (2019). Här framkommer otydlighet och dålig redovisning av struktur kring terapeutkompetens och drar ned betyg till dåligt på delskalor kopplade till detta område. Erfarenhet och träning hos terapeuter i studierna (15) var oftast också begränsad till terapeuter under utbildning, vilket gav små möjligheter att skatta någon av studierna utifrån högsta kvalitet.

### ***Kontroll av följsamhet till behandlingen***

Vid kontroll av följsamhet av själva behandlingen, nyttjades det i samtliga studier systematisk uppföljning, handledning och i vissa fall inspelning av behandlingssessioner för att

öka just följsamhet till manual. Det högsta kravet om ett frekvent användande av en detaljerad skattningsskala vid varje handledningstillfälle för att mer objektivt mäta terapeuters följsamhet, klarade endast Kampmann et al (2016). Övriga studier fick *hygglig* skattning. Samma grundkrav gällande mer specifik kontroll av kompetens vid utförande av interventioner (17) klarade endast Kampmann et al (2016) och Anderson et al., (2013) med nivå *bra* som resultat av frekvent och detaljerad bedömning med hjälp av skattningsskala. Övriga studier saknade skattning, vilket drog ned betyg till *hyggligt*.

### ***Bortfall och statistiska analyser***

Hanteringsstrategi av bortfall (19) från studier redovisas väl och statistiska analyser utgår från "Intent to treat" och räknar alltså med även prematura avhopp från behandlingsbetingelser. De enda studier som inte uppnår betyget bra på denna delskala är Bouchard et al (2017) och Gujjar et al. (2019) som förvisso anger Intent-to-treat men saknar detaljerad beskrivning av konsekvens av detta i analysdel.

### ***Kvantativ översikt av medelvärden***

Utifrån den kvalitativa genomgången och *samtliga artiklars* adderade poäng för de olika delskalorna redovisade i Tabell 3, analyserades och extraherades medelvärden för varje delskala. Syftet var att bättre kunna identifiera eventuella svagheter och brister inom fältet vi granskat genom att översiktligt summera. För att jämföra resultat från liknande studie av Morin et al (2015) räknade vi ut medelvärde för hela skalan  $M=1.36$  ( $SD=0.37$ ). Se Tabell 4 för sammanfattning av medelvärden.

#### **Tabell 4**

Medelvärden (med SD inom parentes) efter kvalitetsbedömning för respektive delskala

Delskalor	M (SD) <sup>a</sup>	Delskalor	M (SD)
1. Beskrivning av sample	1.57 (0.29)	12. Mätpunkter	1,43 (0.53)
2. Kronicitet hos störningen	1.14 (0.14)	13. Behandlingsmanual	1.71 (0.49)
3. Representativitet	1.14 (0.14)	14. Antal terapeuter	1.23 (0.76)
4. Reliabilitet för diagnosen	1.29 (0.24)	15. Terapeutträning	0.86 (0.38)
5. Specifika utfallsmått	1.86 (0.14)	16. Följsamhet till manual	1.00 (0.58)
6. Reliabila utfallsmått	1.86 (0.14)	17. Terapeutkompetens	1.14 (0.69)
7. Blindade bedömare	0.86 (0.48)	18. Samtidiga behandlingar	1.00 (0)
8. Träning av bedömare	1.00 (0.58)	19. Handhavande av bortfall	1.57 (0.79)



9. Fördelning till behandling	1,43 (0.38)	20. Statistiska analyser	2.00 (0)
10. Design	1.86 (0.38)	21. Klinisk signifikans	1.29 (0.49)
11. Statistisk power	0.86 (0.69)	22. Terapitimmor	1.86 (0.38)

<sup>a</sup> Medelvärde för hela skalan är 1.36 (0.37)

De delskalor som hade avvikande, och mest relevant för forskningskvaliteten *lägst* medelvärden, summerades enligt följande kriterier använt av Morin et al (2015) i deras studie för mer riktad kvalitetsgranskning: medelvärden *lägre eller lika med betyget måttligt*. Detta motsvarade 27% av delskalorna i vår studie och gav oss specifikt ett resultat för de i resultatdelen framkomna svagare delskalorna att fokusera på i en vidare diskussion. Dessa var i nummerordning delskalor 7, 8, 11, 15, 16 och 18.

### ***Helskalans poäng per artikel***

För att kunna dra slutsats hur studierna stod sig inbördes extraherades även totalpoäng och sorterades utifrån fallande skattning. Se Tabell 5 för illustration kring hur olika studier har klarat vår genomgång.

## **Tabell 5**

### *Helskalans sammanlagda poäng per artikel <sup>a</sup>*

<u>Studie</u>	<u>Totalpoäng</u>
Gujjar et al., 2019	19p
Bouchard et al., 2017	26p
Meyerbroeker et al., 2013	28p
Lindner et al., 2019	31p
Kampmann et al., 2016	34p
Miloff et al., 2019	35p
Anderson et al., 2013	38p

<sup>a</sup> Resultat sorterade utifrån lägst till högsta poäng. Maxpoäng är 44p

I Östs (Ost, 2008) meta-analys jämförs traditionell KBT med ACT och DBT utifrån Psychotherapy rating scale. I artikeln, där också skalan presenteras för första gången, redovisas inga sammantagna medelvärden, likt Morina et al. (2015), för skalan för de olika behandlingsmodellerna vad vi kan utläsa. Analyser är i stället gjorda på delskalenivå för att utvärdera vad som skiljer KBT från DBT och ACT rent forskningskvalitetsmässigt. Vi lyfte ut

medelvärden från de delskalor vi funnit svagare i vår egen genomgång med motsvarande resultat från denna meta-analys för en enkel illustrerande jämförelse. Se Tabell 6 för detaljer.

### **Tabell 6**

*Illustration av svaga delskalor i vår granskning respektive Öst<sup>a</sup>*

Delskala	(Vår) M (SD)	(Öst ) M (SD)
7. Blindade bedömare	1.14 (0.14)	0.58 (0.50)
8. Träning av bedömare	1.29 (0.24)	0.77 (0.82)
11. Statistisk power	0.86 (0.69)	0.38 (0.80)
15. Träning av terapeuter	0.86 (0.38)	1.42 (0.64)
16. Följsamhet till manual	1.00 (0.58)	0.92 (0.80)
18. Samtidiga behandlingar	1.00 (0.0)	1.00 (0.49)

<sup>a</sup> Från Öst (2008) använde vi medelvärden från Psychotherapy rating scale för de KBT-studier som används som referensvärden i denna meta-analys.

Utifrån denna visuella jämförelse ser vi en generell utmaning med nyttjandet av träning och blindande av bedömare liksom statistisk poweranalys. Syftet är inte att uttala oss om några skillnader, utan att konstatera utmaningarna att utgå från strikta kriterier inom psykoterapiforskningen. Det som möjligtvis är viktigt att lyfta ur tabellen ovan kan vara behovet av fokus på träning och uppföljning av terapeuter inom VR-forskningen.

## Diskussion

Utgår vi från (Higgings, 2011) nivåer för biasskattning ser vi sammantaget att de skattade studierna visar låg eller osäker risk för bias. En del av bedömning av risk för bias är att vara domänspecifik och ta höjd för när och hur en risk för bias både är avgörande och möjlig att kontrollera. För psykoterapiforskning ligger svårigheten i att bibehålla blindhet i behandlingstillfällen. De argument för skattning av osäker risk för bias, handlar primärt om svårigheten att åstadkom bibehålla blindning för både försökspersoner och bedömare för behandlingsbetingelser såväl som testning av utfall av interventioner. För VRET specifikt har ju dessutom experimentell kontrollgrupp VR-utrustning, vilket säkerligen påverkar bias. Vår skattning om osäker risk för bias bedömer vi inte som en anledning att diskvalificera kvalitetsaspekten och samtliga studier delar denna risk för bias för detection och performance bias. En studie sticker ut (Meyerbroeker, 2013) då för lite information finns. Detta kan vara bristande rapportering i själva artikeln, men om inte så skall inte brister på denna nivå behöva förekomma annat än vid sämre planering av forskningen. Den enda delskala som blir rödmarkerad med hög risk för bias är i Gujjar et al. (2017) artikel där både beskrivning saknas för blindning av försökspersoner och behandlare. Vi märker vid skattning av denna artikel att randomisering till behandlingsbetingelser överhuvudtaget är dåligt gjort. Vi ställde oss frågan om denna artikel skulle ha inkluderats utifrån våra kriterier, med tanke på otydlig separation mellan experiment och kontrollgrupp.

Totalt sett gör vi utifrån skattning med Risk of bias tool bedömning av 1) god forskningskvalitet överlag med osäker till låg risk för bias *mellan* artiklarna som helhet, samt 2) för delar av två artiklarna osäker till hög risk för bias för en eller flera domäner (Gujjar et al. 2017 och Meyerbroeker, 2013).

Vi genomförde också en översiktlig genomgång av de sju inkluderade artiklarna med *Psychotherapy rating scale* modifierad av Öst (2008), summerade detta i text och tabellform samt analyserade även delskalor med enkel översyn av medelvärden. Det fanns en initial förväntan utifrån våra inklusionskriterier om RCT:er med aktiv kontrollgrupp, att skattning sannolikt skulle hamna högt kvalitetsmässigt i relation till den generella VRET-studien. Vi kan därför börja med att konstatera att samtliga studier förväntades klara minst en genomsnittlig ”hygglig” nivå gällande de flesta av delskalorna av den anledningen. Ett första visuellt intryck av Tabell 2 är att detta stämde bra, med relativt tillfredställande skattningar totalt sett.

Morina et al. (2015) använde liksom vi *Psychotherapy rating scale* för att kvalitetsbedöma 14 VRET-studier för specifika fobier. Deras medelvärde på helskala efter skattning var 0.95 (0.45), vilket är lägre än vår egen summering på 1.36 (0.37). Denna jämförelse är relevant

eftersom Morina et al. (2015) är den enda studien som vi funnit som är så lik i kombination av skattningsskala som användes, diagnosområde samt behandlingsinriktning. De skattade större bredd och fler studier, men hade lösare inklusionskriterier än oss. Fem av de senast publicerade studierna som vi granskat fanns inte heller tillgängliga vid deras publikationsdatum och skulle tänkbart kunnat lyfta deras medelvärden. De gjorde dessvärre inga analyser av delskalor utifrån forskningskvalitet, utan vi får nöja oss med anekdotiskt konstatera något högre genomsnittligt medelvärde i vår litteraturgranskning.

Vi fann enligt resultatdelen och Tabell 4 sex delskalor med lägre skattningar som potentiellt kan påvisa svagheter inom den forskningsgrenen vi närmast oss. Dessa skulle kunna kategoriseras i fyra tänkbara ”kluster” om man slår samman dessa delskalor, nämligen *kvalitativ bedömning* (7 och 8), *statistisk power* (11), *behandlingskvalitet* (15 och 16) samt strikt *kontroll av samtidiga behandlingar* (18). Vi diskuterar dessa nedan.

I studierna vi undersökt har brister i bedömning framför allt handlat om att säkra blindheten hos bedömare genom fortlöpande kalibrerande kontroller av deras medvetenhet av behandlingsbetingelser och säkra bedömmardrift liksom interbedömmarrelabilitet. I mindre studier lär detta vara resurskrävande, svårt att få till och inte omedelbart tillgängligt för ett mindre lab. Vidare har man i flera fall brustit i poweranalys för att finna en tillräcklig samplestorlek för att göra studieresultat forskningsmässigt valida. Det tredje klustret kopplat till terapeuters arbete och följsamhet har vi anledning att vara kritiska till, då främst erfarenhet hos terapeuter verkar brista eller vara suboptimala för att få goda effektstorlekar i de studier vi undersökt. Vi ser gärna att verkligt erfarna kliniker används i studier snarare än snabbutbildade terapeuter som man får intrycket från i stort sett i samtliga studier. Det finns också goda avsikter när det gäller att säkra följsamhet till behandling i de flesta artiklarna, men generellt faller det på Östs (Ost, 2008) krav att använda en mycket detaljerad skattningsrutin i kombination med systematisk handledning, filmhandledning, etcetera och verkar mer vara en krävande strategi som så här långt har saknats men skulle kunna läggas till.

Enligt vår mening är det inte säkert att detta ökar följsamheten i det kliniska arbetet även om forskningsmässigheten hålls mer intakt. Den siste delskalan enligt ovan handlar om kravet om fullständig kontroll av annan pågående behandling, psykologisk och medicinsk, vilken endast en studie klarade av att efterleva. Utmaningen med en så strikt exkludering är förmodligen svårighet att rekrytera med bibehållet representativt sample, eller lyckas rekrytera i tillräcklig utsträckning överhuvudtaget. Samsjuklighet är stor i denna patientgrupp. Det vi inte vet är hur verklighetsnära dessa kvalitetskrav är och hur man lyckas efterleva detta i vedertagen forskning.

## **Summerande diskussion och kritisk reflektion**

Huvudfrågorna kring översiktlig forskningskvalitet gällande VRET för fobier har i viss utsträckning besvarats. Urvalet vi har gjort av studier verkar ha haft god forskningsmässig kvalitet. Utifrån egen skattning och nyanserade diskussion kring de metaanalyser vi funnit tillgängliga, verkar dessa kunna stå sig i jämförelse och indikera god forskningskvalitet hos de inkluderade studierna. Det är dock möjligt att utforska alternativ till förbättring och utveckling både av använd metodik och vår egen ansats.

Vi valde i uppsatsen att inte alls redovisa kvantitativ data och effektstorlekar för de studier vi tagit med, utan hållit oss helt till diskussion kring kvalitativ sådan. Detta är en brist, även om det ger mer utrymme för utforskande av kvalitativ data. Det vi inte heller kan uttala oss om är hur forskningskvaliteten ser ut för området VR-baserad KBT som helhet, eftersom vår selektering inte inkluderat så många nyare studier, pågående utveckling, trender och vilka utmaningar som finns inom detta snabbt växande forskningsfält. Utöver det har vi inte heller fördjupat oss i flera viktiga aspekter kring VR-behandling såsom etiska tillkortakommanden och problem med privata finansiella intressen där framför allt leverantörer av programvara som används vid VRET ofta samarbetar tätt med publicerande forskargrupper och skulle kunna påverka risk of bias.

## Referenser

- Alonso, J., Angermeyer, M. C., Bernert, S., Bruffaerts, R., Brugha, T. S., Bryson, H., de Girolamo, G., Graaf, R., Demyttenaere, K., Gasquet, I., Haro, J. M., Katz, S. J., Kessler, R. C., Kovess, V., Lépine, J. P., Ormel, J., Polidori, G., Russo, L. J., Vilagut, G., Almansa, J., ... ESEMeD/MHEDEA 2000 Investigators, European Study of the Epidemiology of Mental Disorders (ESEMeD) Project (2004). Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta psychiatrica Scandinavica. Supplementum*, (420), 21–27. <https://doi.org/10.1111/j.1600-0047.2004.00327.x>
- Anderson, P. L., Price, M., Edwards, S. M., Obasaju, M. A., Schmertz, S. K., Zimand, E., & Calamaras, M. R. (2013). Virtual reality exposure therapy for social anxiety disorder: a randomized controlled trial. *Journal of consulting and clinical psychology*, 81(5), 751–760. <https://doi.org/10.1037/a0033559>
- Bijl, R. V., van Zessen, G., Ravelli, A., de Rijk, C., & Langendoen, Y. (1998). The Netherlands Mental Health Survey and Incidence Study (NEMESIS): objectives and design. *Social psychiatry and psychiatric epidemiology*, 33(12), 581–586. <https://doi.org/10.1007/s001270050097>
- Bouchard, S., Dumoulin, S., Robillard, G., Guitard, T., Klinger, É., Forget, H., Loranger, C., & Roucaut, F. X. (2017). Virtual reality compared with *in vivo* exposure in the treatment of social anxiety disorder: a three-arm randomised controlled trial. *The British journal of psychiatry : the journal of mental science*, 210(4), 276–283. <https://doi.org/10.1192/bjp.bp.116.184234>
- Choy, Y., Fyer, A. J., & Lipsitz, J. D. (2007). Treatment of specific phobia in adults. *Clinical psychology review*, 27(3), 266–286. <https://doi.org/10.1016/j.cpr.2006.10.002>
- Clark D. M. (1999). Anxiety disorders: why they persist and how to treat them. *Behaviour research and therapy*, 37 Suppl 1, S5–S27. [https://doi.org/10.1016/s0005-7967\(99\)00048-0](https://doi.org/10.1016/s0005-7967(99)00048-0)
- Craske, M. G., Niles, A. N., Burklund, L. J., Wolitzky-Taylor, K. B., Vilardaga, J. C., Arch, J. J., Saxbe, D. E., & Lieberman, M. D. (2014). Randomized controlled trial of cognitive behavioral therapy and acceptance and commitment therapy for social phobia: outcomes and moderators. *Journal of consulting and clinical psychology*, 82(6), 1034–1048. <https://doi.org/10.1037/a0037212>

- Davidson, J. R., Foa, E. B., Huppert, J. D., Keefe, F. J., Franklin, M. E., Compton, J. S., Zhao, N., Connor, K. M., Lynch, T. R., & Gadde, K. M. (2004). Fluoxetine, comprehensive cognitive behavioral therapy, and placebo in generalized social phobia. *Archives of general psychiatry*, *61*(10), 1005–1013.  
<https://doi.org/10.1001/archpsyc.61.10.1005>
- Deacon, B. J., Farrell, N. R., Kemp, J. J., Dixon, L. J., Sy, J. T., Zhang, A. R., & McGrath, P. B. (2013). Assessing therapist reservations about exposure therapy for anxiety disorders: the Therapist Beliefs about Exposure Scale. *Journal of anxiety disorders*, *27*(8), 772–780.  
<https://doi.org/10.1016/j.janxdis.2013.04.006>
- Dellazizzo, L., Potvin, S., Luigi, M., & Dumais, A. (2020). Evidence on Virtual Reality-Based Therapies for Psychiatric Disorders: Meta-Review of Meta-Analyses. *Journal of medical Internet research*, *22*(8), e20889.  
<https://doi.org/10.2196/20889>
- Difede, J., & Hoffman, H. G. (2002). Virtual reality exposure therapy for World Trade Center Post-traumatic Stress Disorder: a case report. *Cyberpsychology & behavior : the impact of the Internet, multimedia and virtual reality on behavior and society*, *5*(6), 529–535.  
<https://doi.org/10.1089/109493102321018169>
- Fernández-Alvarez J, Colombo D, Botella C, García-Palacios A, Riva G. (2019). Virtual reality for anxiety and stress-related disorders: a SWOT analysis. In: Cipresso P, Serino S, Villani D, editors. *Pervasive Computing Paradigms for Mental Health*. New York, USA: Springer; 2019.
- Fournier, J. C., DeRubeis, R. J., Hollon, S. D., Dimidjian, S., Amsterdam, J. D., Shelton, R. C., & Fawcett, J. (2010). Antidepressant drug effects and depression severity: a patient-level meta-analysis. *JAMA*, *303*(1), 47–53.  
<https://doi.org/10.1001/jama.2009.1943>
- GBD 2016 Disease and Injury Incidence and Prevalence Collaborators (2017). Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet (London, England)*, *390*(10100), 1211–1259. [https://doi.org/10.1016/S0140-6736\(17\)32154-2](https://doi.org/10.1016/S0140-6736(17)32154-2)

- Cuijpers, P., van Straten, A., Bohlmeijer, E., Hollon, S. D., & Andersson, G. (2010). The effects of psychotherapy for adult depression are overestimated: a meta-analysis of study quality and effect size. *Psychological medicine*, *40*(2), 211–223. <https://doi.org/10.1017/S0033291709006114>
- Gujjar, K. R., van Wijk, A., Kumar, R., & de Jongh, A. (2019). Efficacy of virtual reality exposure therapy for the treatment of dental phobia in adults: A randomized controlled trial. *Journal of anxiety disorders*, *62*, 100–108. <https://doi.org/10.1016/j.janxdis.2018.12.001>
- Gustavsson, A., Svensson, M., Jacobi, F., Allgulander, C., Alonso, J., Beghi, E., Dodel, R., Ekman, M., Faravelli, C., Fratiglioni, L., Gannon, B., Jones, D. H., Jennum, P., Jordanova, A., Jönsson, L., Karampampa, K., Knapp, M., Kobelt, G., Kurth, T., Lieb, R., ... CDBE2010Study Group (2011). Cost of disorders of the brain in Europe 2010. *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology*, *21*(10), 718–779. <https://doi.org/10.1016/j.euroneuro.2011.08.008>
- Harvey, A. G., & Gumpert, N. B. (2015). Evidence-based psychological treatments for mental disorders: modifiable barriers to access and possible solutions. *Behaviour research and therapy*, *68*, 1–12. <https://doi.org/10.1016/j.brat.2015.02.004>
- Higgins, J. P., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., Savovic, J., Schulz, K. F., Weeks, L., Sterne, J. A., Cochrane Bias Methods Group, & Cochrane Statistical Methods Group (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ (Clinical research ed.)*, *343*, d5928. <https://doi.org/10.1136/bmj.d5928>
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of consulting and clinical psychology*, *59*(1), 12–19. <https://doi.org/10.1037//0022-006x.59.1.12>
- Kampmann, I. L., Emmelkamp, P. M., Hartanto, D., Brinkman, W. P., Zijlstra, B. J., & Morina, N. (2016). Exposure to virtual social interactions in the treatment of social anxiety disorder: A randomized controlled trial. *Behaviour research and therapy*, *77*, 147–156. <https://doi.org/10.1016/j.brat.2015.12.016>
- LeBeau, R. T., Glenn, D., Liao, B., Wittchen, H. U., Beesdo-Baum, K., Ollendick, T., & Craske, M. G. (2010). Specific phobia: a review of DSM-IV specific phobia



- and preliminary recommendations for DSM-V. *Depression and anxiety*, 27(2), 148–167. <https://doi.org/10.1002/da.20655>
- Lindner, P., Miloff, A., Fagnäs, S., Andersen, J., Sigeman, M., Andersson, G., Furmark, T., & Carlbring, P. (2019). Therapist-led and self-led one-session virtual reality exposure therapy for public speaking anxiety with consumer hardware and software: A randomized controlled trial. *Journal of anxiety disorders*, 61, 45–54. <https://doi.org/10.1016/j.janxdis.2018.07.003>
- Mavranouzouli, I., Mayo-Wilson, E., Dias, S., Kew, K., Clark, D. M., Ades, A. E., & Pilling, S. (2015). The Cost Effectiveness of Psychological and Pharmacological Interventions for Social Anxiety Disorder: A Model-Based Economic Analysis. *PloS one*, 10(10), e0140704. <https://doi.org/10.1371/journal.pone.0140704>
- McCann, R. A., Armstrong, C. M., Skopp, N. A., Edwards-Stewart, A., Smolenski, D. J., June, J. D., Metzger-Abamukong, M., & Reger, G. M. (2014). Virtual reality exposure therapy for the treatment of anxiety disorders: an evaluation of research quality. *Journal of anxiety disorders*, 28(6), 625–631. <https://doi.org/10.1016/j.janxdis.2014.05.010>
- Meyerbroeker, K., Morina, N., Kerkhof, G. A., & Emmelkamp, P. M. (2013). Virtual reality exposure therapy does not provide any additional value in agoraphobic patients: a randomized controlled trial. *Psychotherapy and psychosomatics*, 82(3), 170–176. <https://doi.org/10.1159/000342715>
- Miloff, A., Lindner, P., Dafgård, P., Deak, S., Garke, M., Hamilton, W., Heinsoo, J., Kristoffersson, G., Rafi, J., Sindemark, K., Sjölund, J., Zenger, M., Reuterskiöld, L., Andersson, G., & Carlbring, P. (2019). Automated virtual reality exposure therapy for spider phobia vs. in-vivo one-session treatment: A randomized non-inferiority trial. *Behaviour research and therapy*, 118, 130–140. <https://doi.org/10.1016/j.brat.2019.04.004>
- Morina, N., Ijntema, H., Meyerbröker, K., & Emmelkamp, P. M. (2015). Can virtual reality exposure therapy gains be generalized to real-life? A meta-analysis of studies applying behavioral assessments. *Behaviour research and therapy*, 74, 18–24. <https://doi.org/10.1016/j.brat.2015.08.010>
- Oosterink, F. M., de Jongh, A., & Aartman, I. H. (2008). What are people afraid of during dental treatment? Anxiety-provoking capacity of 67 stimuli

- characteristic of the dental setting. *European journal of oral sciences*, 116(1), 44–51. <https://doi.org/10.1111/j.1600-0722.2007.00500.x>
- Oosterink, F. M., de Jongh, A., & Hoogstraten, J. (2009). Prevalence of dental fear and phobia relative to other fear and phobia subtypes. *European journal of oral sciences*, 117(2), 135–143. <https://doi.org/10.1111/j.1600-0722.2008.00602.x>
- Opriş, D., Pinteă, S., García-Palacios, A., Botella, C., Szamosközi, Ş., & David, D. (2012). Virtual reality exposure therapy in anxiety disorders: a quantitative meta-analysis. *Depression and anxiety*, 29(2), 85–93. <https://doi.org/10.1002/da.20910>
- Parsons, T. D., & Rizzo, A. A. (2008). Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: a meta-analysis. *Journal of behavior therapy and experimental psychiatry*, 39(3), 250–261. <https://doi.org/10.1016/j.jbtep.2007.07.007>
- Piaggio, G., Elbourne, D. R., Pocock, S. J., Evans, S. J., Altman, D. G., & CONSORT Group (2012). Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. *JAMA*, 308(24), 2594–2604. <https://doi.org/10.1001/jama.2012.87802>
- Rothbaum, B. O., Anderson, P., Zimand, E., Hodges, L., Lang, D., & Wilson, J. (2006). Virtual reality exposure therapy and standard (in vivo) exposure therapy in the treatment of fear of flying. *Behavior therapy*, 37(1), 80–90. <https://doi.org/10.1016/j.beth.2005.04.004>
- Shuster, J. J. (2011). Review of Cochrane handbook for systematic reviews for interventions, Version 5.10. *Research Synthesis Methods*, 2(2), 126–130. <https://doi-org.ludwig.lub.lu.se/10.1002/jrsm.38>
- Stagnaró, J. C., Cía, A. H., Aguilar Gaxiola, S., Vázquez, N., Sustas, S., Benjet, C., & Kessler, R. C. (2018). Twelve-month prevalence rates of mental disorders and service use in the Argentinean Study of Mental Health Epidemiology. *Social psychiatry and psychiatric epidemiology*, 53(2), 121–129. <https://doi.org/10.1007/s00127-017-1475-9>
- Stein, D. J., Lim, C., Roest, A. M., de Jonge, P., Aguilar-Gaxiola, S., Al-Hamzawi, A., Alonso, J., Benjet, C., Bromet, E. J., Bruffaerts, R., de Girolamo, G., Florescu, S., Gureje, O., Haro, J. M., Harris, M. G., He, Y., Hinkov, H., Horiguchi, I., Hu, C., Karam, A., ... WHO World Mental Health Survey Collaborators (2017). The cross-national epidemiology of social anxiety disorder: Data from

- the World Mental Health Survey Initiative. *BMC medicine*, 15(1), 143.  
<https://doi.org/10.1186/s12916-017-0889-2>
- Stein, J. P., & Ohler, P. (2017). Venturing into the uncanny valley of mind-The influence of mind attribution on the acceptance of human-like characters in a virtual reality setting. *Cognition*, 160, 43–50.  
<https://doi.org/10.1016/j.cognition.2016.12.010>
- Swift JK, Greenberg RP. (2014). A treatment by disorder meta-analysis of dropout from psychotherapy. *J Psychother Integr* 2014;24(3):193-207. [doi: 10.1037/a0037512]
- Världshälsoorganisationen. (1992). International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10). Den svenska titeln är Internationell statistisk klassifikation av sjukdomar och relaterade hälsoproblem – Systematisk förteckning, svensk version 2016 (ICD-10-SE). World Health Organization 1992
- Wardenaar, K. J., Wanders, R., Ten Have, M., de Graaf, R., & de Jonge, P. (2017). Using a hybrid subtyping model to capture patterns and dimensionality of depressive and anxiety symptomatology in the general population. *Journal of affective disorders*, 215, 125–134. <https://doi.org/10.1016/j.jad.2017.03.038>
- Wechsler, T. F., Kümpers, F., & Mühlberger, A. (2019). Inferiority or Even Superiority of Virtual Reality Exposure Therapy in Phobias?-A Systematic Review and Quantitative Meta-Analysis on Randomized Controlled Trials Specifically Comparing the Efficacy of Virtual Reality Exposure to Gold Standard *in vivo* Exposure in Agoraphobia, Specific Phobia, and Social Phobia. *Frontiers in psychology*, 10, 1758. <https://doi.org/10.3389/fpsyg.2019.01758>
- Wolitzky-Taylor, K. B., Horowitz, J. D., Powers, M. B., & Telch, M. J. (2008). Psychological approaches in the treatment of specific phobias: a meta-analysis. *Clinical psychology review*, 28(6), 1021–1037.  
<https://doi.org/10.1016/j.cpr.2008.02.007>
- Zlomke, K., & Davis, T. E., 3rd (2008). One-session treatment of specific phobias: a detailed description and review of treatment efficacy. *Behavior therapy*, 39(3), 207–223. <https://doi.org/10.1016/j.beth.2007.07.003>
- Ost L. G. (2008). Efficacy of the third wave of behavioral therapies: a systematic review and meta-analysis. *Behaviour research and therapy*, 46(3), 296–321.  
<https://doi.org/10.1016/j.brat.2007.12.005>