TECHNICAL COMMENTARY

NeuRA Discover. Conquer. Cure.

SCHIZOPHRENIA LIBRARY

Virtual Reality and Avatar therapy

Introduction

Virtual reality is a modern computerised real-time technology using graphics, sounds and other sensory input, which creates an interactive computer-mediated world. Virtual reality applications have been primarily used for the assessment and treatment of anxieties and phobias. More recently they have been used to examine the perception of emotion and the emotional responses of people with schizophrenia during simulated social encounters, with the aim of improving social skills, cognitive functioning, and treatment adherence.

Avatar therapy uses virtual reality to recreate the faces and voices of hallucinations. Using a computer program, patients create an avatar with the help of a therapist. The avatar's voice is selected to match the voice heard by the patient. While the patient establishes dialogues with the avatar, the therapist manages the avatar so that it is gradually controlled by the patient. The avatar's mode changes from persecutory to supportive during the therapy sessions. These sessions can be recorded for the patient to take home.

Method

We have included only systematic reviews (systematic literature search, detailed methodology with inclusion/exclusion criteria) published in full text, in English, from the year 2000 that report results separately for people with a diagnosis of schizophrenia, schizoaffective disorder, schizophreniform disorder or first episode schizophrenia. Reviews were identified by searching the databases MEDLINE, EMBASE, CINAHL, Current Contents, PsycINFO and the Cochrane library. Hand searching reference lists of identified reviews was also conducted. When multiple copies of reviews were found, only the most recent version was included. Reviews with pooled data are prioritised for inclusion.

Review reporting assessment was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist which describes a preferred way to present a meta-analysis[1]. Reviews with less than 50% of items checked have been excluded from the library. The PRISMA flow diagram is a suggested way of providing information about studies included and excluded with reasons for exclusion. Where no flow diagram has been presented by individual reviews, but identified studies have been described in the text, reviews have been checked for this item. Note that early reviews may have been guided by less stringent reporting checklists than the PRISMA, and that some reviews may have been limited by journal guidelines.

Evidence was graded using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group approach where high quality evidence such as that gained from randomised controlled trials (RCTs) may be downgraded to moderate or low if review and study quality is limited, if there is inconsistency in results, indirect comparisons, imprecise or sparse data and high probability of reporting bias. It may also be downgraded if risks associated with the intervention or other matter under review are high. Conversely, low quality evidence such as that gained from observational studies may be upgraded if effect sizes are large or if there is a dose dependent response. We have also taken into account sample size and whether results are consistent, precise and direct with low associated risks (see end of table for an explanation of these terms)[2]. The resulting table represents an objective summary of the available evidence, although the conclusions are solely the opinion of staff of NeuRA (Neuroscience Research Australia).

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

NeuRA Discover. Conquer. Cure.

SCHIZOPHRENIA LIBRARY

Virtual Reality and Avatar therapy

Results

We found two systematic reviews that met our inclusion criteria[3, 4].

- Moderate to low quality evidence finds avatar therapy may improve general symptoms of schizophrenia and attitudes towards voices. There were no consistent benefits for acceptance of voices, quality of life, or treatment retention. Review authors suggest that the therapy needs further testing in large, well-designed, and well-reported RCTs.
- Moderate to low quality evidence finds no benefits of virtual reality training for cognition, social skills, acceptability of treatment or study retention.

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Aali G, Kariotis T, Shokraneh F

Avatar Therapy for people with schizophrenia or related disorders

Cochrane Database of Systematic Reviews 2020; 5: Cd011898

View review abstract online

Comparison 1	**Avatar therapy plus standard care vs. standard care alone.**
Summary of evidence	**Low quality evidence (very small samples, imprecise, direct) finds no consistent benefits of avatar therapy for symptoms, relapse or treatment retention.**
Mental state	
Better general symptoms on the PSYRATS with avatar therapy; 1 RCT, N = 19, MD = -5.51, 95%CI -9.15 to -1.87, *p* = 0.003 *No significant differences on the PANSS;* Positive symptoms: 1 RCT, N = 19, MD = -1.93, 95%CI -5.10 to 1.24, *p* = 0.23 General symptoms: 1 RCT, N = 19, MD = -8.07, 95%CI -19.49 to 3.35, *p* = 0.17	
Relapse	
No significant differences in; Rehospitalization: 1 RCT, N = 19, RD = 0.00, 95%CI -0.20 to 0.20, *p* = 1.00 Needing for counselling: 1 RCT, N = 19, RR = 1.85, 95%CI 0.09 to 40.05, *p* = 0.70	
Leaving the study early	
No significant difference; 1 RCT, N = 26, RR = 11.27, 95%CI 0.70 to 181.41, *p* = 0.09	
Comparison 2	**Avatar therapy plus standard care vs. supportive therapy plus standard care.**
Summary of evidence	**Moderate to low quality evidence (small to medium-sized samples, imprecise, direct) finds avatar therapy may improve general symptoms more than supportive therapy, with no consistent benefits for attitudes or acceptance of voices, quality of life, or treatment retention.**

Mental state	
Better general symptoms on the PSYRATS with avatar therapy; 1 RCT, N = 124, MD = -4.74, 95%CI -8.01 to -1.47, *p* = 0.005	
Attitudes towards and acceptance of voices	
Better attitude towards voices with avatar therapy; 1 RCT, N = 124, MD = -8.39, 95%CI -14.31 to -2.47, *p* = 0.005 *Better acceptance of voices with supportive therapy;* 1 RCT, N = 124, MD = 4.73, 95%CI 1.40 to 8.06, *p* = 0.005 *No significant difference in the power of the voices;* 1 RCT, N = 115, MD = -0.36, 95%CI -0.89 to 0.17, *p* = 0.18	
Quality of life	
No significant differences in quality of life as measured by; Manchester Short Assessment of Quality of Life: 1 RCT, N = 120, MD = 2.69, 95%CI -1.48 to 6.86, *p* = 0.21 Rosenberg Self-Esteem Scale: 1 RCT, N = 121, MD = 0.55, 95%CI -1.55 to 2.65, *p* = 0.61	
Leaving the study early	
No significant difference in rates of leaving the study early for any reason; 1 RCT, N = 150, RR = 1.06, 95%CI 0.59 to 1.89, *p* = 0.85	
Risks	No differences in anxiety symptoms.
Consistency in results‡	Not applicable; all analysis included 1 RCT.
Precision in results§	Imprecise
Directness of results‖	Direct

Välimäki M, Hätönen HM, Lahti ME, Kurki M, Hottinen A, Metsäranta K, Riihimäki T, Adams CE

Virtual reality for treatment compliance for people with serious mental

illness	
Cochrane Database of Systematic Reviews 2014; Issue 10. Art. No.: CD009928. DOI: 10.1002/14651858.CD009928.pub2 View review abstract online	
Comparison	**Virtual reality training (5 to 12 weeks) vs. standard care.**
Summary of evidence	**Moderate to low quality evidence (small to medium-sized samples, unable to assess precision, direct) finds no benefits of virtual reality training for cognition, social skills, acceptability of treatment or study retention.**
Cognitive functioning	
No significant differences between groups; 1 RCT, N = 27, MD = 4.67, 95%CI -1.76 to 11.10, $p > 0.05$ Authors rate this evidence as low quality.	
Social skills	
No significant differences between groups; 1 RCT, N = 64, MD = -2.30, 95%CI -8.13 to 3.53, $p > 0.05$ Authors rate this evidence as low quality.	
Acceptability of treatment	
No significant differences between groups; 2 RCTs, N = 92, RD = 0.05, 95%CI -0.09 to 0.19, $p > 0.05$, $I^2 = 0\%$ Authors rate this evidence as low quality.	
Loss to follow-up	
No significant differences between groups; 3 RCTs, N = 156, RD = 0.02, 95%CI -0.08 to 0.12, $p > 0.05$, $I^2 = 0\%$ Authors rate this evidence as low quality.	
Consistency in results	Consistent where applicable (> 1 RCT).
Precision in results	Unable to assess, no measure of precision is reported.
Directness of results	Direct

Explanation of acronyms

CI = confidence interval, I^2 = the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance), MD = mean difference, N = number of participants, *p* = statistical probability of obtaining that result ($p < 0.05$ generally regarded as significant), PANSS = Positive And Negative Syndrome Scale, RCT = randomized controlled trial, PSYRATS = Psychotic Symptom Rating Scale, RD = risk difference, RR = risk ratio, vs. = versus

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Explanation of technical terms

* Bias has the potential to affect reviews of both RCT and observational studies. Forms of bias include; reporting bias – selective reporting of results; publication bias - trials that are not formally published tend to show less effect than published trials, further if there are statistically significant differences between groups in a trial, these trial results tend to get published before those of trials without significant differences; language bias – only including English language reports; funding bias - source of funding for the primary research with selective reporting of results within primary studies; outcome variable selection bias; database bias - including reports from some databases and not others; citation bias - preferential citation of authors. Trials can also be subject to bias when evaluators are not blind to treatment condition and selection bias of participants if trial samples are small[5].

† Different effect measures are reported by different reviews.

Prevalence refers to how many existing cases there are at a particular point in time. Incidence refers to how many new cases there are per population in a specified time period. Incidence is usually reported as the number of new cases per 100,000 people per year. Alternatively some studies present the number of new cases that have accumulated over several years against a person-years denominator. This denominator is the sum of individual units of time that the persons in the population are at risk of becoming a case. It takes into account the size of the underlying population sample and its age structure over the duration of observation.

Reliability and validity refers to how accurate the instrument is. Sensitivity is the proportion of actual positives that are correctly identified (100% sensitivity = correct identification of all actual positives) and specificity is the proportion of negatives that are correctly identified (100% specificity = not identifying anyone as positive if they are truly not).

Mean difference scores refer to mean differences between treatment and comparison groups after treatment (or occasionally pre to post treatment) and in a randomised trial there is an assumption that both groups are comparable on this measure prior to treatment. Standardised mean differences are divided by the pooled standard deviation (or the standard deviation of one group when groups are homogenous), which allows results from different scales to be combined and compared. Each study's mean difference is then given a weighting depending on the size of the sample and the variability in the data. Less than 0.4 represents a small effect, around 0.5 a medium effect, and over 0.8 represents a large effect[5].

Relative risk (RR) refers to the probability of a reduction (< 1) or an increase (> 1) in a particular outcome in a treatment group, or a group exposed to a risk factor, relative to the comparison group. For example, a RR of 0.75 translates to a reduction in risk of an outcome of 25% relative to those not receiving the treatment or not exposed to the risk factor. Conversely, a RR of 1.25 translates to an increased risk of 25% relative to those not receiving treatment or not having been exposed to a risk factor. A RR of 1.00 means there is no difference between groups. A medium effect is considered if RR > 2 or < 0.5 and a large effect if RR > 5 or < 0.2[6]. Odds ratios (ORs) are similar to RRs, but they are based on the probability of an event occurring divided by the probability of that event not

occurring. ORs and RRs are similar in size when the event is rare, such as with schizophrenia. lnOR stands for logarithmic OR where a lnOR of 0 shows no difference between groups. Hazard ratios (HRs) measure the effect of an explanatory variable on the hazard or risk of an event.

Correlation coefficients (eg, r) indicate the strength of association or relationship between variables. They can provide an indirect indication of prediction, but do not confirm causality due to possible and often unforseen confounding variables. An r of 0.10 represents a weak association, 0.25 a medium association and 0.40 and over represents a strong association. Unstandardised (*b*) regression coefficients indicate the average change in the dependent variable associated with a 1 unit change in the independent variable, statistically controlling for the other independent variables. Standardised regression coefficients represent the change being in units of standard deviations to allow comparison across different scales.

‡ Inconsistency refers to differing estimates of effect across studies (i.e. heterogeneity or variability in results) that is not explained by subgroup analyses and therefore reduces confidence in the effect estimate. I^2 is the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance) - 0% to 40%: heterogeneity might not be important, 30% to 60%: may represent moderate heterogeneity, 50% to 90%: may represent considerable heterogeneity and over this is considerable heterogeneity. I^2 can be calculated from Q (chi-square) for the test of heterogeneity with the following formula[5];

$$I^2 = \left(\frac{Q - df}{Q}\right) \times 100\%$$

§ Imprecision refers to wide confidence intervals indicating a lack of confidence in the effect estimate. Based on GRADE recommendations, a result for continuous data (standardised mean differences, not weighted mean differences) is considered imprecise if the upper or lower confidence limit crosses an effect size of 0.5 in either direction, and for binary and correlation data, an effect size of 0.25. GRADE also recommends downgrading the evidence when sample size is smaller than 300 (for binary data) and 400 (for continuous data), although for some topics, these criteria should be relaxed[7].

‖ Indirectness of comparison occurs when a comparison of intervention A versus B is not available but A was compared with C and B was compared with C that allows indirect comparisons of the magnitude of effect of A versus B. Indirectness of population, comparator and/or outcome can also occur when the available evidence regarding a particular population, intervention, comparator, or outcome is not available and is therefore inferred from available evidence. These inferred treatment effect sizes are of lower quality than those gained from head-to-head comparisons of A and B.

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

References

1. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMAGroup (2009): Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *British Medical Journal* 151: 264-9.
2. GRADEWorkingGroup (2004): Grading quality of evidence and strength of recommendations. *British Medical Journal* 328: 1490.
3. Valimaki M, Hatonen HM, Lahti ME, Kurki M, Hottinen A, Metsaranta K*, et al.* (2014): Virtual reality for treatment compliance for people with serious mental illness. *Cochrane Database of Systematic Reviews* 10: CD009928.
4. Aali G, Kariotis T, Shokraneh F (2020): Avatar Therapy for people with schizophrenia or related disorders. *Cochrane Database of Systematic Reviews* 5: Cd011898.
5. CochraneCollaboration (2008): Cochrane Handbook for Systematic Reviews of Interventions. Accessed 24/06/2011.
6. Rosenthal JA (1996): Qualitative Descriptors of Strength of Association and Effect Size. *Journal of Social Service Research* 21: 37-59.
7. GRADEpro (2008): [Computer program]. Jan Brozek, Andrew Oxman, Holger Schünemann. *Version 32 for Windows*