| **Section and Topic** | **Item #** | **Checklist item** | Location where item is reported | **Supporting information** |
| --- | --- | --- | --- | --- |
| **TITLE** | | |  |  |
| Title | 1 | Identify the report as a systematic review. | Line 1 |  |
| **ABSTRACT** | | |  |  |
| Abstract | 2 | See the PRISMA 2020 for Abstracts checklist. | Line 8 |  |
| **INTRODUCTION** | | |  |  |
| Rationale | 3 | Describe the rationale for the review in the context of existing knowledge. | Section 2, Line 297-298 |  |
| Objectives | 4 | Provide an explicit statement of the objective(s) or question(s) the review addresses. | Section 4, Line 334-338 |  |
| **METHODS** | | |  |  |
| Eligibility criteria | 5 | Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses. | Section 4, Line 353-359 |  |
| Information sources | 6 | Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted. | Section 4, Line 359-362 |  |
| Search strategy | 7 | Present the full search strategies for all databases, registers and websites, including any filters and limits used. | Section 4, Line 342-352 |  |
| Selection process | 8 | Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process. | Section 4, Line 362-371 |  |
| Data collection process | 9 | Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process. | Section 4, Line 369-371 |  |
| Data items | 10a | List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect. | Section 5, Line 373-376  Section 5, Line 421-431 | The eligible outcomes for the systematic review were categorised as:  “*Study type (T) was categorised as either:*   1. *Non-comparative (N) – these are studies which demonstrated a BET undertaken only in a CAVE.* 2. *Technology comparative (CT) – these are studies which compared a BET undertaken in a CAVE with those undertaken in other VE(s).* 3. *Reality comparative (CR) – these are studies which compared a BET undertaken in a CAVE with those undertaken. These studies may also compare with other virtual environment(s).*”   Any study which met one of these outcomes was included. All studies were read by the corresponding author and no automation methods or tools were used to collect data from the studies.  For each outcome 1-3, the following data was collected:   * Study type (T). * Brief overview of study purpose. * Author(s). * Year of publication. * Country of publication. |
| 10b | List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information. | Section 4, Line 371  (“See Data Availability Statement” - this document) | For each outcome 1-3, any study which demonstrated “*psychophysiological or behavioural analyses of CAVE participants whilst undertaking BETs*”, had the following data extracted:   * Variable of interest. * Experimental measure. |
| Study risk of bias assessment | 11 | Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process. | Section 4, Line 371  (“See Data Availability Statement” - this document) | As the systematic review aimed to understand what the purpose of the study was and whether the study was comparative or not, the actual study results were of less interest.  More information was collected from the psychophysiological and behavioural studies, however, the data extracted was only the variable(s) of interest and the experimental measures. The results of the experiment(s) are not considered for any study in this review, as the review aims instead to qualitatively categorise the purpose, type and comparativeness of previous work.  The authors therefore see no domains of bias for the data extracted (study type (T), brief overview of study purpose, author(s), date of publication and country of publication).  The only possible domain of bias for the data collected in the review is from those psychophysiological and behavioural studies, where the domain of bias is:   * The measurement technique used.   The authors recognise that bias could be present in the measurement technique used in the study, if the study funders provided the product, technique, or questionnaire, or if the author(s) were motivated to use their own measurement technique.  The measurement techniques used in these studies are presented in Table 3, and the most prevalent techniques (by number of occurrences) are discussed. Therefore, the actual performance, sensitivity or benefits of the individual measurement techniques is not considered, meaning any bias in the studies towards one technique is not carried over into the results of the systematic review. Thus, this authors treat this bias domain as negligible. |
| Effect measures | 12 | Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results. | Section 4, Line 371  (“See Data Availability Statement” - this document) | The eligible outcomes for the systematic review were categorised as:  “*Study type (T) was categorised as either:*   1. *Non-comparative (N) – these are studies which demonstrated a BET undertaken only in a CAVE.* 2. *Technology comparative (CT) – these are studies which compared a BET undertaken in a CAVE with those undertaken in other VE(s).* 3. *Reality comparative (CR) – these are studies which compared a BET undertaken in a CAVE with those undertaken. These studies may also compare with other virtual environment(s).*”   For each outcome, the following measure(s) were used in the synthesis or presentation of results:   * Study type (T). * Brief overview of study purpose. * Author(s). * Year of publication. * Country of publication. |
| Synthesis methods | 13a | Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)). | Section 4, Line 371  (“See Data Availability Statement” - this document) | As all studies were included in the first synthesis, there was no process to decide which studies to include or exclude. From the full screening, studies which measured a psychophysiological or behavioural variable were included in the second synthesis. |
| 13b | Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions. | Section 4, Line 371  (“See Data Availability Statement” - this document) | All studies were read in full by the corresponding author and a brief summary of study purpose was written. The studies were then categorised using the thematic analysis technique mentioned in Section 5 Lines 386-388, and the experimental procedure was analysed to understand if a CAVE, other virtual environment, or real-world equivalent setting had been used. This data was used to identify the study type (T). Additionally, the date of publication, author names, country of publication were extracted from the publisher page or DOI link. From the full screening, studies which measured a psychophysiological or behavioural variable were included in the second synthesis. Here, the variable(s) of interest (dependent variables) were extracted along with their experimental measurement technique. |
| 13c | Describe any methods used to tabulate or visually display results of individual studies and syntheses. | Section 4, Line 371  (“See Data Availability Statement” - this document) | No particular methods or techniques were used to tabulate the results of the first synthesis. The study category, description, and type (T) were placed into a table and alphabetised by lead author surname. The studies were group by their category for ease of presentation.  All studies included in the review were then presented by Year of Publication (Figure 6), Country of Publication (Figure 7, and Thematic Analysis (Figure 9) using Microsoft Excel. All studies included in the review were presented in a Co-Authorship Network (Figure 10) by exporting all references from EndNote 20 into .RIS format, and using VOSviewer to visualise the results. The network was weighted by link strength and year of publication data was used for visualisation purposes.  For the second synthesis (psychophysiological and behavioural studies), the data collected on Variable of Interest and Experimental Measure were tabulated for each study using Microsoft Excel. The total number of studies for each Variable of Interest was displayed at the bottom of the table and the total number of studies which used each Experimental Measure was extracted and discussed in Section 7. |
| 13d | Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used. | Section 4, Line 371  (“See Data Availability Statement” - this document) | The first synthesis methodology was to simply group studies together by their data. This data has already been described as:   * Study type (T). * Brief overview of study purpose. * Author(s). * Year of publication. * Country of publication.   Simple summation was used to synthesise the results and present the studies from a scale of most to least common for each data type.  The second synthesis used the same technique but specifically for psychophysiological and behavioural Variables of Interest and their corresponding Experimental Measures.  As previously stated, the systematic review did not investigate the findings of the individual studies. Rather, it aimed to identify the experimental setup, procedure and measures of each study. Therefore, no synthesis was done on the experimental results. |
| 13e | Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression). | Section 4, Line 371  (“See Data Availability Statement” - this document) | Simple summation was used to synthesise the results and present the studies from a scale of most to least common for each data type. The causes of heterogeneity among study results was not explored as it is not possible in the context of the systematic review. |
| 13f | Describe any sensitivity analyses conducted to assess robustness of the synthesized results. | Section 4, Line 371  (“See Data Availability Statement” - this document) | Sensitivity analysis was not suitable for the systematic review, as the review retrieved and presented objective data collected from the studies (author, year, study type, study purpose). The results of the systematic review directly reflect the contents of the studies and meta-analysis was not undertaken. |
| Reporting bias assessment | 14 | Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases). | Section 4, Line 371  (“See Data Availability Statement” - this document) | Risk of bias was not assessed as no data was missing from the data collection criteria for all studies. |
| Certainty assessment | 15 | Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome. | Section 4, Line 371  (“See Data Availability Statement” - this document) | Certainty or confidence for all outcomes was not assessed due to the objective nature of the data extracted from studies. Results were synthesised presenting the most to least common year of publication, country of publication, study category, psychophysiological and behavioural variable of interest and measure. The certainty of the review in its scope is discussed directly in Section 6.3. |
| **RESULTS** | | |  |  |
| Study selection | 16a | Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram. | Section 4, Line 366-368 |  |
| 16b | Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded. | Section 4, Line 371  (“See Data Availability Statement” – Systematic Review Data.xlsx) | See “Systematic Review Data.xlsx”. Reason for study exclusion is mentioned for a number of studies which would appear to meet the inclusion criteria. |
| Study characteristics | 17 | Cite each included study and present its characteristics. | Section 5, Line 390-392 |  |
| Risk of bias in studies | 18 | Present assessments of risk of bias for each included study. | Section 4, Line 371  (“See Data Availability Statement” - this document) | As mentioned in Item 11, the risk of bias for the systematic review is minimal and has been described for all studies, rather than each individual included study. |
| Results of individual studies | 19 | For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots. | Table 2, Table 3  Figure 6, 7, 8 and 9 | The precision of results is not applicable to the results of this systematic review due to the objective nature of the data extracted. |
| Results of syntheses | 20a | For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies. | Section 4, Line 371  (“See Data Availability Statement” - this document) | As mentioned in Item 11, the risk of bias for the systematic review is minimal and has been described for all studies, rather than each individual included study. |
| 20b | Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect. | Section 4, Line 371  (“See Data Availability Statement” - this document) | Example 1: In a review examining the effects of aspirin for primary prevention of cardiovascular disease, the authors report for a meta-analysis of risk ratios the number of included studies and participants, summary estimate and its 95% confidence interval, the I 2 measure of inconsistency, and they translate the relative effect into absolute terms: “Twelve studies [each study cited], including a total of 159,086 patients, reported on the rate of major bleeding complications. Aspirin use was associated with a 46% relative risk increase of major bleeding complications (risk ratio 1.46; 95% CI, 1.30-1.64;)  Example 2: In a review examining the effect of exercise programmes for ankylosing spondylitis, the authors report for a metaanalysis of mean differences the number of included studies and participants, summary estimate and its 95% confidence interval, the I2 measure of inconsistency, and they translate the absolute effect into relative terms and describe the clinical importance of the result: “Physical function (BASFI, 0 to 10 scale; lower score indicates higher function): Seven studies (312 participants) found a reduction in physical function score with exercise versus no intervention at the end of the intervention (mean difference (MD) -1.3, 95% confidence interval (CI) -1.7 to -0.9); absolute risk difference 13% (95% CI 9% to 17%); relative change 32% (95% CI 23% to 42%); Analysis 1.1). The statistical heterogeneity was not important (I²= 23%). There was no important clinical meaningful benefit.” (72) |
| 20c | Present results of all investigations of possible causes of heterogeneity among study results. | Section 4, Line 371  (“See Data Availability Statement” - this document) | Simple summation was used to synthesise the results and present the studies from a scale of most to least common for each data type. The causes of heterogeneity among study results was not explored as it is not possible in the context of the systematic review. |
| 20d | Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results. | Section 4, Line 371  (“See Data Availability Statement” - this document) | Sensitivity analysis was not suitable for the systematic review, as the review retrieved and presented objective data collected from the studies (author, year, study type, study purpose). The results of the systematic review directly reflect the contents of the studies and meta-analysis was not undertaken. |
| Reporting biases | 21 | Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed. | Section 4, Line 371  (“See Data Availability Statement” - this document) | Risk of bias was not assessed as no data was missing from the data collection criteria for all studies. |
| Certainty of evidence | 22 | Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed. | Section 4, Line 371  (“See Data Availability Statement” - this document) | Certainty or confidence for all outcomes was not assessed due to the objective nature of the data extracted from studies. Results were synthesised presenting the most to least common year of publication, country of publication, study category, psychophysiological and behavioural variable of interest and measure. The certainty of the review in its scope is discussed directly in Section 6.3. |
| **DISCUSSION** | | |  |  |
| Discussion | 23a | Provide a general interpretation of the results in the context of other evidence. | Section 6.1, Section 6.2 |  |
| 23b | Discuss any limitations of the evidence included in the review. |  | The limitations of the evidence included in the systematic review are as follows:   * Inclusion criteria.   + The systematic review required studies to be in the context of “built environment tasks”, however, this is a difficult context to specify. The systematic review attempted to address this by specifically excluding certain applications (see Section 4 Line 357-359). * Applications of CAVEs outside of literature.   + The systematic review investigated the use of virtual environments (such as CAVEs) for replicating real-world tasks. However, it is possible that research and applications of CAVE technology in this area has not been published or has been undertaken by private companies.   As the systematic review is concerned only with the characteristics of the study conducted rather than the results of the individual studies, the limitations of the (objective) evidence included in the review are minimal. |
| 23c | Discuss any limitations of the review processes used. | Section 6.3 |  |
| 23d | Discuss implications of the results for practice, policy, and future research. | Section 7, Section 9 |  |
| **OTHER INFORMATION** | | |  |  |
| Registration and protocol | 24a | Provide registration information for the review, including register name and registration number, or state that the review was not registered. | Review not registered. |  |
| 24b | Indicate where the review protocol can be accessed, or state that a protocol was not prepared. | Section 4. |  |
| 24c | Describe and explain any amendments to information provided at registration or in the protocol. | N/A. |  |
| Support | 25 | Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review. | The Ph.D. programme of the corresponding author is co-sponsored by the University of Birmingham (50%) and Fulcro Engineering Services (50%). |  |
| Competing interests | 26 | Declare any competing interests of review authors. | N/A |  |
| Availability of data, code and other materials | 27 | Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review. | Section 4, Line 371 (“See Data Availability Statement”) |  |

*From:*  Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71