Learning from comics versus non-comics material in education: Systematic review and meta-analysis

Marianna Pagkratidou[1], Neil Cohn[2], Phivos Phylactou[3,4], Marietta Papadatou-Pastou[5] , Gavin Duffy[6]

[1]School of Psychology, Dublin City University, Dublin, Ireland

[2] Department of Communication and Cognition, Tilburg University, The Netherlands

[3] School of Physical Therapy, University of Western Ontario, London, Ontario, Canada

[4] The Gray Centre for Mobility and Activity, Parkwood Institute of Rehabilitation, London, Ontario, Canada

[5] Department of Primary Education, National and Kapodistrian University of Athens, Athens, Greece

[6] School of Electrical and Electronic Engineering, TU Dublin, Ireland

Author note

Marianna Pagkratidou https://orcid.org/0000-0002-0161-1554

Neil Cohn https://orcid.org/0000-0003-0074-1995

Phivos Phylactou https://orcid.org/0000-0002-7333-8761

Marietta Papadatou-Pastou https://orcid.org/0000-0002-2834-4003

Gavin Duffy https://orcid.org/0000-0003-3049-7030

Authors declare no conflicts of interest.

Preprint version: 10 May 2024 (v3).

This preprint is under peer-review.

Funding

This project is supported by the School of Electrical and Electronic Engineering, Technological University Dublin through the SellSTEM (Spatially Enhanced Learning Linked to STEM) Marie Skłodowska-Curie Innovative Training Network funding - European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 956124.

.

Abstract

The past decades have seen a growing use of comics (i.e., sequential presentation of images and/or text) educational material. However, there are inconsistent reports regarding their effectiveness. In this study, we aim to systematically review empirical studies that have investigated the use of comics in education; and to quantitatively explore these effects using a meta-analysis. To do so, we will search PubMed, Scopus, and Web of Science for studies employing an experimental design that uses comics education material compared to non-comics education material in general population samples. ~~We hypothesize that if learning via comics operates in the same way as learning via non-comics material, namely texts), then, comics and texts will have the same impact on learning; however, if the combination of text and images confers greater communicability beyond learning via texts, then we hypothesize that comics will have a greater impact on learning.~~ Our findings will not only shed light on whether comics are equally or more effective education material than texts, but also on the conditions in which comics can foster learning.

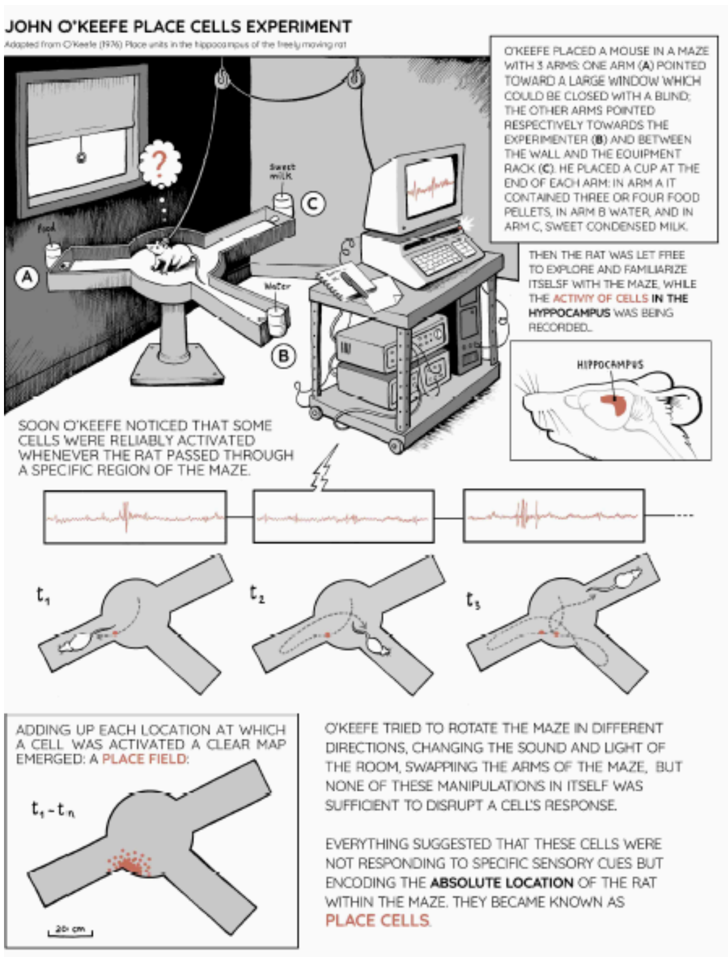
Keywords: comics cognition, visual language, visual language fluency, spatial cognition, STEM learning

Introduction

Imagine you are a teacher eager to introduce your class to the seminal work of O'Keefe and Nadel (1978; Nobel Prize, 2014) on place cells. While numerous educational materials convey this information, they often present students with extensive text pushing them to infer the designs, perhaps with loosely associated diagrams. As an alternative, consider the comic in Figure 1, where this information is integrated into a coherent discourse. Have you ever considered utilizing comics for education?

Figure 1

Comics (Bach et al., 2018; ©Matteo Farinella) on the Place cells experiment (O'Keefe & Nadel, 1978)



The past decades have witnessed a growing use of and advocacy for comics in education settings (Cary 2004; Farinella, 2020; Sousanis, 2015; Topkaya et al., 2023), with school librarians indicating that the inclusion of comic books attract more students to use school libraries (Lo et al., 2022); and with more than 75% of educators[1] using comics in their classrooms daily either as independent reading or as a supplement to the main lesson or as their main teaching track (Comic Book Legal Defense Fund, 2019). However, there are inconsistent findings regarding the effectiveness of learning when using comics compared to non-comcis material, namely texts. In this study, we will conduct a systematic review on using comics in education, as well as a meta-analysis to quantify the overall effect of empirical studies that used comics versus texts. Furthermore, we will explore whether comics affect learning differently when it comes to Science, Technology, Engineering and Mathematics (STEM) compared to non-STEM fields.

Comics can be defined as a particular type of social object, used by people of a particular cultural orientation, which use visual language (sequential images) and writing, typically associated with contexts and styles (Chute, 2008; Cohn, 2012; Cohn & Magliano, 2019). However, this definition is complicated by the term "comics" in education being conflated by how people use the term in the first place (Cohn, 2013a; Gavaler, 2022). In one sense, "comics in education" is meant as the promotion of the structural properties of the "medium" (i.e., sequential text/image units). Such an advantage would manifest in educational materials being created in this manner. For example, this sense may include cases where textbooks or educational materials are created using multimodal text-image units put into a sequence.

[1] Note that the educators were surveyed by the Comic Book Legal Defense Fund

Another sense is that comics as existing social objects are beneficial in education. In these cases, the published works that carry the designation of "comics" are used in educational contexts. This would include using published comics to teach about literature or to teach second languages, or using Art Speigelman's memoir *Maus* about the holocaust as a way to teach history. This does not necessarily use the structural properties of visual and multimodal expression as a means to educate (such as in comparison to standard textbook formats), but rather uses published literature that may otherwise be viewed as entertainment within educational contexts. While advocacy for "comics in education" often conflates these senses both in theory and practice, these distinctions are important for disentangling their purported advantages.

Indeed, although often associated with entertainment, the comics "medium" has been shown to be valuable and beneficial for conducting and communicating science and they have been further used for educational purposes for over 80 years by a range of other scholars (Farinella, 2018; McCloud, 1993; Topkaya et al., 2023; Yang, 2008). In the realm of conducting science, comics have been used as stimuli for studying areas such as the theory of mind (Baron-Cohen, Leslie, & Frith, 1986) in individuals with Autism Spectrum Disorder (ASD; Sivaratnam et al., 2012), and spatial mental representations (Pagkratidou et al., 2023). In the realm of communicating science, comics are utilized in both STEM educational contexts, such as health-related education (Green & Myers, 2010), science communication (Aleixo & Sumner 2017; Bach et al., 2017; Farinella, 2018), and non-STEM educational contexts, such as law (Botes, 2017), and business education (Short et al., 2013).

Comics as a "medium" typically use a combination of pictures and text, and while these forms are often thought to be different, recent research has shown that the sequential images used in comics and texts share similarities in two levels: structural similarities and similarities in how they are processed (Cohn, 2012; 2013; Cohn & Kutas, 2017; Loschky et al., 2018; Loschky et al., 2019). In terms of structural similarities, the graphics of comics are created using a "visual language" with its own lexicon and grammar structured in a similar manner as sentences in texts (Cohn 2013; Cohn & Kutas, 2017). In terms of processing, reading comics involves the mechanisms of retrieval, integration, and updating, as does reading texts (Cohn 2020; Hagoort, 2003; Van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998).

Despite these similarities in structure and processing, an important difference between comics and purely textual documents is that comics are visualized in a sequence of images (Cohn, 2018), whereas no visualization is offered in the case of texts. This sequence of images in the panels contains more information than the individual units of text, such as character dialogues, objects, feelings, thoughts conveyed in word or thought bubbles, stars and symbols (2012). Background information provides the reader with a combination of linguistic information and visual presentations in richer examples and in a more engaging way that could lead to effective communication (White-Schwoch & Rapp, 2011; Cohn, 2012). Whether the additional information presented in the background of comics panels confers a benefit to foster learning remains an open question.

A recent meta-analysis by Topkaya et al (2023) on the effectiveness of using comics in education along with many others have argued for the educational benefit of comics (e.g., Damopolii et al., 2021; Farinella, 2018; Leiner et al., 2018; Muzumdar et al., 2018), with several reasons offered for such an advantage. First, the integration of text and pictures provides readers with multimodal units that package information in a more effective way than the independent associations made in typical textbooks (Cohn, 2013b; Mayer, 2002; Schnotz & Rasch, 2005). The sequential nature of these units also provides a coherent and guided structure for conveying complex information. Their pictorial nature may also seem less imposing than pure text, while many educational comics have content that appears more engaging or entertaining, making them more accessible. All of these reasons have been offered as the benefits of comics for education, yet disentangling any true benefit and its reasons remains underexplored.

Building on these questions, recent studies have compared comics with non-comics material in a variety of formal and non-formal educational settings. For example, Leiner et al. (2018) asked patients to study health-related instructions presented as a comic versus instructions presented as text alone, and then to complete a questionnaire about their understanding of the instructions; results showed that understanding and recalling was significantly better via the comics than via the texts. However, when Lin et al. (2015) conducted a mixed-methods quasi experimental study to examine the public knowledge of nanotechnology, when using comics versus texts as education material, there was no significant difference between the two media. Interestingly, when Lestari and Mustadi (2020) conducted a quasi-experimental study with a pretest-post-test design to examine the storytelling skills of fifth-grade students using animated video versus comics, they found that animated videos were more effective than comics. Overall, there is no consensus on whether comics are effective to improve learning (Cohn, 2020; Farinella, 2020).

The lack of consensus in the findings of studies that have investigated the effectiveness of comics when it comes to learning could be attributed to various differences in the experimental procedures. These differences include the target educational fields of *STEM* (e.g., Wiseman et al., 2021; Shimazaki et al., 2021; Freedman et al., 2022) versus *non-STEM* (e.g., Cohen et al., 2020; Istiq'faroh et al., 2020; Owens et al., 2020), or target population of *students* (e.g., Lin et al., 2016; Damopolii et al., 2021) versus the *public* (e.g., Rodriguez et al., 2016; Shimazaki et al., 2021; Wiseman et al., 2021). The media comparisons also differ between *text* (e.g., Leiner et al., 2018), *animation* (e.g., Walsh et al., 2021), or *video* (e.g.., Kolberg et al., 2021). In addition, several studies show moderation by individual differences, such as *visual language fluency* (e.g., Cohn, 2013; Spiegel et al., 2013; Wang et al., 2019), *student's achievement level* (e.g., Lin et al., 2016; Topkaya et al., 2020); while others don't (e.g., Damopolii et al., 2021). Finally, the intervention style differs across the studies, with some studies following a *quasi experimental design* (e.g., Kim et al., 2017), others an *experimental design* (e.g., Muzumdar et al., 2018), some including a *follow up measurement* (e.g., Sumengen et al., *2023)* others not (e.g., Shimazaki et al., 2021), and some utilizing comics as a *main* (e.g., Albero et al., 2021) while others as a *complementary* (e.g., Khoii et al., 2010) *intervention*.

Given the conflicting results, the main question driving the current research is whether comics, compared to non-comics education material, namely text material, are an effective learning tool[2]. To our knowledge, no past research has integrated the range of empirical studies that conducted interventions to improve learning using comics versus non-comics materials. The aim of the present study is to systematically review and to quantify using meta-analysis the overall effect of comics vs texts material that have been used in empirical studies that targeted learning for STEM and non-STEM fields. We specifically aim to examine potential differences between STEM and non-STEM contexts, as panels in comics, by providing external visualizations to students, might scaffold learning differently for STEM and non-STEM topics, especially for individuals with low spatial abilities, considering the relationship between visualization, spatial ability efficacy, and STEM learning (Newcombe, 2013; 2017; Zhu et al., 2023). To do so, we will measure factors related

[2] Note that we will measure which medium is more effective for learning by calculating the effect sizes of the knowledge measurement and by comparing that among the studies that used comics vs texts.

to the experimental process, the individual differences, and the intervention style. Our main questions are the following:

Table 1.

Study design plan

	Research question	Hypotheses	Sampling plan	Analysis plan	Interpretation given different outcomes	Theory that could be shown wrong by the outcomes
1.	What are the claimed benefits of comics vs text for education?	We refrain from forming concrete hypotheses, as our analysis will be exploratory in nature.	Systematic Review	Meta-analysis	This is an exploratory study. Interpretation for all results will be presented in the discussion.	This is an exploratory study. Explanations for all results will be presented in the discussion.
2.	Is there a difference in the putative effectiveness of comics in STEM vs non-STEM subjects?	We refrain from forming concrete hypotheses, as our analysis will be exploratory in nature.	Systematic Review		This is an exploratory study. Interpretation for all results will be presented in the discussion.	This is an exploratory study. Explanations for all results will be presented in the discussion.
3.	Is there a moderating effect in the putative relationship between comics and learning of factors such as age, target population, experimental design, intervention type and alternative non-educational material such as videos etc?	We refrain from forming concrete hypotheses, as our analysis will be exploratory in nature.	Systematic Review	Meta-analysis	N/A	This is an exploratory study. Explanations for all results will be presented in the discussion.

Method

To report the systematic review and meta-analyses we will follow the guidelines of the PRISMA statement (Page et al., 2020); and we will present the PRISMA 2020 Main Checklist and the PRISMA 2020 Abstract Checklist in the Supplementary Material. All codes, data and materials will be shared on this OSF repository: https://osf.io/cmqb6/?view_only=98677e4fcab84e47968556c7958817f3

Study search

Our search strategy will aim for completeness. The studies that we will include in the meta-analyses are going to be located via a multi-step process:

(i) *Via databases:* We will search the electronic databases *PubMed (accessed through https://pubmed.ncbi.nlm.nih.gov/)*, *Scopus (accessed through scopus.com)*, Google Scholar (accessed through https://scholar.google.com/), *Open Grey* (accessed through opengrey.eu) and *Web of Science (accessed through EBSCOhost)* in "All Fields" from their inception. The search terms that we will use are the following:

Table 2.

Literature search terms. To be included, studies will need to refer to at least one term from each column (i.e., [item from Comics education material] AND [item from Learning]).

Comics education material	Learning
comic*	education*
cartoon	knowledge
graphic novel	learn*

(ii) *Via other methods:* We will also hand-search the reference lists of review papers we might retrieve via our search strategy. Moreover, we will request missing data and unpublished datasets from the corresponding authors of each included paper (or of the papers for which inclusion will be contingent on obtaining data that will not have been reported in the manuscript) using the email address provided in the article and in case this is no longer in use, we will search online for a more recent one. The strategy of contacting the authors will be by sending a reminder after 10 days and another one 5 more days.

Study selection

We will use the open-source online reference management software-platforms that will help usin the screening of big datasets. Specifically, we will use Zotero v.6.0.23 (Roy Rosenzweig Center for History and New Media, 2021) to create a database of all retrieved records and to identify and merge duplicates and Covidence (https://www.covidence.org/) to conduct the title and abstract screening, the full paper screening, and to export and report the records. We will sequentially screen titles, abstracts, and full-texts. All materials will be made available, including zotero bib files, and csv files from Covidence where reviews stay archived in their website. Authors MP and PP will select the studies independently and author GD will resolve any inconsistencies (i) by taking the records to the full-text stage of the review in the titles and abstracts screening stages, even if only one author accepts them or is unsure and (ii) via discussion in the full-text stage. MPP author, specializing in meta-analyses, will resolve any cases of uncertainty. Studies excluded at the full-text stage together with the reasons for exclusion will be listed in a Table included in the Supplementary Material.

Inclusion and exclusion criteria

The following criteria will be set for inclusion of an individual study in the meta-analyses:

1. Participants: We will include studies that measured general population samples of both adults and children - public and students.

2. Sufficient data: We will use data that are either reported in the studies in a usable way for the analysis or provided by the authors after request. To be usable the data should enable us to calculate an effect size (Cohen's d or Cohen's d_{rm}) that can be used to pool an overall effect size in a random effects meta-analysis model (e.g., mean, *sd/SE, n*; see *Statistical Analyses*). If no data are provided by the studies or the authors and if possible, we will extract data from figures with the use of WebPlotDigitizer (https://automeris.io/WebPlotDigitizer/).

3. Comics as education material: We will include studies that operationally defined comics as "whatever people call comics"; and if there is no clear definition of what comics is, then we will include studies that operationally defined comics as the use of sequential images and writing in objects identified as "comics" (Cohn, 2020).

4. Publication type: We will include only empirical studies that compare comics with any texts material; and we will exclude any case studies of individuals, review studies, and meta-analyses.

5. Research design: We will include only studies that include an experimental and a control group.

6. *Outcome variable*: We will include only studies that have an outcome variable for any kind of "knowledge" measurement.

Of note, publication language will not be taken into consideration as we will not only include reports written in the languages spoken by the research team, but also we will include papers in any language that can be translated through ChatGPT4 (OpenAI, 2023; https://openai.com/gpt-4) large language model, using the prompt ’Translate the following text to English:’ followed by the section text in quotation marks.”

Data extraction

MP and PP will perform the data extraction independently. Should any inconsistencies arise, those will be resolved through discussion. A third reviewer, MPP, will resolve any remaining ambiguities. Inter-rater reliability will be calculated among the two reviewers, using Cohen's Kappa estimate (see Conry-Murray et al., 2023). Specifically, Cohen’s Kappa will be initially calculated for a small pool of papers (10%). If agreement is not satisfactory (i.e., <80%), the two reviewers will meet to discuss disagreements. The process will be repeated until satisfactory agreement is reached (i.e., >80%) or until screening is complete. We will not replace any missing data. The following data will be extracted from each study:

1. *Sample size:* We will extract the number of participants in the experimental and control groups.
2. *Year of publication:* We will extract the year of publication of the study and enter it into the database numerically.
3. *Location:* We will extract the country where the study was conducted.
4. *Education field*: Depending on the learning material used, we will categorize the studies as STEM (code=1) and non-STEM (code=2). STEM will include a broad spectrum of STEM fields such as Science, Technology, Engineering, Mathematics, Energy, Medicine, Biology, Space and Geography studies (and any other related disciplines; and non-STEM will include all the fields that are not related with STEM).

5. *Sex of the participants:* We will extract the sex of the participants separately for each group (experimental vs. control) as male, female, or other (if reported), for descriptive purposes
6. *Mean age of the participants:* We will extract the mean age of the participants in each group and the average age between the groups will be entered into the database numerically. In the case that only an age range is reported (e.g., 7-10-year-old), then the middle of the range will be recorded.
7. *Type of participants:* We will code the target group of the educational intervention as pupils (=1), University/College students (=2) and general public (=3)
8. *Visual Language Reading Fluency:* We will extract data regarding whether studies have collected data on some type of individual differences measurement, such as the Visual Language Reading Fluency (Cohn, 2014), by using a yes (=1) / no (=2) coding. Also, we will extract the level of Visual Language Reading Fluency by coding it to either low (=1) or high (=2).
9. *Intervention type*: We will extract information regarding the type of intervention by using the coding system of comics as main teaching material (=1) - referring to studies that used comics[3] and text as the main medium to educate the participants - and comics as supplementary material (=2) - referring to studies that used comics and text as a complementary medium to the existing course or text material.
10. *Duration of the intervention*: We will document the duration of the intervention in days, to examine the duration of the learning effect during the intervention period.

[3] Note that we are interested in studies that use comics as one group and text as another group. We will code a control group as a control group if it serves the purpose of a control group, whether it is being labeled or not. Also, in the case a study results in more than one effect size, we will explore how this might bias our results by employing a multi-model meta-analysis, by including a third-level representing the study.

11. *Achievement level*: We will extract information about whether the level of knowledge-achievement in the topic of interest has been taken into consideration in the studies by using a yes (=1) / no (=2) coding. In addition, we will extract the level of knowledge-achievement as specified by the studies, by coding the level of the participants as low (=1), medium (=2), and high (=3).
12. *Treatment condition (comics)*: We will document information about whether the researchers have created their own comics (=1) or have used existing ones (=2).
13. *Control condition (texts)*: We will document information regarding the type of the text; and for any non-comic educational material used, we will document exploratory the information by using the following coding system: text (=1), photo (=2), animation video (=3), games (=4), etc.
14. *Spatial cognition*: We will extract whether data on spatial cognition is reported by using a yes (=1) / no (=2) coding. In addition, we will extract the level of spatial cognition as specified by the studies, by coding the level of the participants as low (=1), medium (=2), and high (=3).
15. *Experimental design:* We will code whether studies have used a pre- and post- design by using a yes (=1) / no (=2) coding.
16. *Experimental type*: We will code whether the studies conducted a true experiment (=1) or a quasi experiment (=2).
17. *Follow up measurement*: We will extract, for descriptive purposes, as a categorical variable whether or not a follow-up study has been conducted, to examine the duration of the learning effect after the intervention period.

Statistical analysis

For the statistical analyses, we will use R (v. 4.2.1 for macOS) and RStudio (2022.07.1 Build 554; R Core Team 2021) with the "meta" (v. 5.5-0; Balduzzi et al., 2019), "metafor" (v. 3.4-0; Viechtbauer, 2010), "dmetar" (v.0.0.9000; Harrer et al., 2019), and "tidyverse" (v. 1.3.2; Wickham et al., 2019) packages (or later versions). MP and PP will perform the analysis and MPP will assess for reliability. Using the esc library (Lüdecke, 2018) we will estimate the effect size of each study, and we will calculate Cohen's *d* (Cohen, 1988). Specifically, effect sizes will be calculated as the standardized difference in knowledge between individuals in the comics condition and a control condition. As such, a *d* equal to zero will correspond to no difference between the comics and the control conditions, values greater than zero will indicate an improvement in learning due to the use of comics, and values lower than zero will indicate a decrease in learning due to the use of comics. ~~For our meta-analysis, we consider an effect size $d = 0.4$ as the minimum effect of interest. This is derived by a previous meta-analysis that investigated the effectiveness of comics in education (Topkaya et al, 2023). In the study by Topkaya et al. (2023), a meta-analysis with subject area as a moderator variable (similar to our STEM and non-STEM categorisation) resulted in an overall effect size of $g = 0.50$, 95% CI [.33, .68].~~ Using the lower bound of the confidence intervals as a heuristic, and considering the Hedge's *g* correction, we define $d = 0.4$ as the minimum effect of interest to guide our analyses and interpretation.

We will pool the effect sizes using a random effects model (Fleiss, 1993) using the inverse variance method, where variance includes both within- and between-study variance. We will use the random effects model, which controls for the possibility that the true effect size may vary from study to study, providing a more flexible and robust analysis (Kanters, 2021). If the total number of identified studies is less than $k = 20$, we will apply the Knapp-Hartung (Knapp & Hartung, 2003) adjustment to our random effects model. The Knapp-Hartung adjustment has been shown to reduce the chance of false positive findings

(Langan et al., 2018), and has been recommended when the number of studies is less than 20 (IntHout et al., 2014). Our hypotheses will be tested at an $\alpha = 0.05$. In cases where $p < \alpha$, we will use equivalence testing to investigate whether the null hypothesis (i.e., no difference) can be accepted (see Lakens et al., 2020). Specifically, if the 90% confidence intervals of the overall effect size includes the smallest difference of interest within their bounds (d = -0.2 and d = 0.2), the null hypothesis will be accepted. In addition, we will report prediction intervals, which refer to the range of effects that future studies are expected to reach based on the given evidence. We will estimate the maximum-likelihood using tau^2 and the Q-profile method using the tau and tau^2 confidence intervals. Heterogeneity will be quantified using the I^2 index (Higgins et al., 2003) and will be visually examined through p-value drapery plots (Rücker & Schwarzer, 2021). Mikolajewicz and Komarova (2019) provide a comprehensive summary for how Cohen's d, tau^2, Q, and I^2 are formulated. If heterogeneity is suspected, as reflected through significant heterogeneity (I^2 , $p < .05$), high heterogeneity variance ($I^2 > 25\%$; Higgins et al., 2003), or prediction regions broader than the overall p-curve in the drapery plot (Rücker & Schwarzer, 2021), then moderator analyses will be conducted. To proceed with moderator analyses, we will require at least five data points per level of each categorical variable, or at least five data points (i.e., five studies) per continuous variable. We will use forest plots for data visualization. Further, we will employ multiple approaches to investigate potential heterogeneity and small study bias considering that it is recommended to employ various methods to investigate bias for meta-analyses in psychology (Carter et al., 2019). Specifically, we will use p-value drapery plots (Rücker & Schwarzer, 2021) and funnel plots to visually investigate small study bias. In addition, p-curves will be generated (Simonsohn et al., 2020) and tested for skewness and flatness using a χ^2 Binomial test, while the symmetry of the effect sizes will be examined using the Egger's regression test (Egger et al., 1997). If small study bias is identified through the visual inspection and/or the Egger's regression test, we will proceed with adjustments of the funnel plot using the Duval and Tweedie (2000) trim and fill method. The adjusted funnel plot will then be visually inspected to identify the direction of bias, and an adjusted overall effect size based on the trim and fill correction will be estimated.

If we find five or more studies with pre-post design we will conduct a three-level meta-analysis. This pre-post design meta-analysis will include a third level (study level), corresponding to the independent study by which each effect size will be estimated, to account for the additional variance introduced by the violation of the independence of the included effect sizes (Van den Noortgate et al., 2013). Effect sizes for the three-level meta-analysis will be estimated as repeated-measures Cohen's *d* (d_{rm}), corresponding to the standardized difference of knowledge before and after the comics intervention. If a correlation coefficient is not provided by the primary studies to calculate d_{rm}, a correlation coefficient (*r*) will be approximated using the formula $r = (t^2 * (sd^2_{\text{pre}} + sd^2_{\text{post}}) - N * mean^2_{\text{change}}) / (2 * t^2 * sd_{\text{pre}} * sd_{\text{post}})$, where *t* is the corresponding *t*-value, sd_{pre} and sd_{post} is the standard deviation of the sample before and after the intervention, respectively, *N* is the sample size, and $mean_{\text{change}}$ is the difference in means before and after the intervention.

As for the sensitivity analysis, and in order to evaluate the stability of the population-effect size we will follow Westerhausen and Papadatou-Pastou (2022); if one study has a weight in the analysis of 25% or above, we will repeat the meta-analysis without this study. In addition, we will conduct outlier identification and influential cases identification. For the influential cases identification we will produce a Baujat plot, an overall influence diagnostic plot, and two leave-one-out meta-analysis plots (one sorted by effect size and the other by I2) for each meta-analysis. Regarding the continuous data meta-analysis, prior to the influential cases identification analysis, we will recalculate the overall effect using the Paule-Mandel method as the T2 estimator. Finally, as for the internal validity (risk of bias) of included studies we will use Cochrane's risk of bias tool to examine the relevant domain-based risk of bias, and we will present the results of this assessment in a graphical format using the "robvis" package (McGuinness & Higgins 2020; McGuinness 2019).

References

Ferrer-Albero, C., & Díez-Domingo, J. (2021). Does a comic style informed assent form improve comprehension for minors participating in clinical trials?. *Clinical Ethics, 16*(1), 37-45. DOI: https://doi.org/10.1177/1477750920930377

Aleixo, P. A., & Sumner, K. (2017). Memory for biopsychology material presented in comic book format. *Journal of Graphic Novels and Comics, 8*(1), 79-88. DOI: https://doi.org/10.1080/21504857.2016.1219957

Botes, M. (2017). Using comics to communicate legal contract cancellation. *The Comics Grid: Journal of Comics Scholarship, 7*(1), 14. DOI: https://doi.org/10.16995/cg.100

Bach, B., Wang, Z., Farinella, M., Murray-Rust, D., & Henry Riche, N. (2018, April). Design patterns for data comics. *In Proceedings of the 2018 chi conference on human factors in computing systems* (pp. 1-12). DOI: http://dx.doi.org/10.1145/3173574.3173612

Baron‐Cohen, S., Leslie, A. M., & Frith, U. (1986). Mechanical, behavioural and intentional understanding of picture stories in autistic children. *British Journal of developmental psychology, 4*(2), 113-125. DOI: https://doi.org/10.1111/j.2044-835X.1986.tb01003.x

Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). How to perform a meta-analysis with R: a practical tutorial. *BMJ Ment Health, 22*(4), 153-160. DOI: https://doi.org/10.1136/ebmental-2019-300117

Carter, E. C., Schönbrodt, F. D., Gervais, W. M., & Hilgard, J. (2019). Correcting for bias in psychology: A comparison of meta-analytic methods. *Advances in Methods and Practices in Psychological Science*, *2*(2), 115-144. https://doi.org/10.1177/2515245919847196

Cary, S. (2004). Going graphic: *Comics at work in the multilingual classroom* (Vol. 102). Portsmouth, NH: Heinemann.

C. B. L. D. F. (2019). *CBLDF's Comics In Education Survey Reveals Top Comics in Schools & A Broad Embrace of the Format*. Comic Book Legal Defense Fund. Retrieved October 14, 2023, from https://cbldf.org/2019/10/cbldfs-comics-in-education-survey-reveals-top-comics-in-schools-a-broad-embrace-of-the-format/

Chute, H. (2008). Comics as literature? Reading graphic narrative. *Pmla, 123*(2), 452-465. DOI: https://doi.org/10.1632/pmla.2008.123.2.452

Cohen, J. (1988). Set correlation and contingency tables. *Applied psychological measurement, 12*(4), 425-434. DOI: https://doi.org/10.1177/014662168801200410

Cohen, E. L., Wasserman, J. A., Schlue, L. M., Keely, C., & Russell, A. (2020). Seeing is believing: The role of imagery fluency in narrative persuasion through a graphic novel. *Psychology of Popular Media, 9*(2), 176–183. DOI: https://doi.org/10.1037/ppm0000218

Cohn, N. (2012). Comics, linguistics, and visual language: The past and future of a field. In Linguistics and the Study of Comics (pp. 92-118). *London: Palgrave Macmillan UK.* DOI: https://doi.org/10.1057/9781137004109_5

Cohn, N. (2013a). Navigating comics: An empirical and theoretical approach to strategies of reading comic page layouts. *Frontiers in psychology, 4,* 186. DOI: https://doi.org/10.3389/fpsyg.2013.00186

Cohn, N. (2013b). Beyond speech balloons and thought bubbles: The integration of text and image. *Semiotica*. 2013(197): 35-63.

Cohn, N. (2014). The Visual Language Fluency Index: A measure of "comic reading expertise." *Visual Language Lab: Resources*. www.visuallanguagelab.com/resources.html. First posted online April 2014.

Cohn, N. (2018). In defense of a "grammar" in the visual language of comics. *Journal of Pragmatics, 127*, 1-19. DOI: https://doi.org/10.1016/j.pragma.2018.01.002

Cohn, N., & Kutas, M. (2017). What's your neural function, visual narrative conjunction? Grammar, meaning, and fluency in sequential image processing. *Cognitive research: principles and implications, 2*(1), 1-13. DOI: https://doi.org/10.1186/s41235-017-0064-5

Cohn, N., & Magliano, J. P. (2020). Editors' Introduction and Review: Visual Narrative Research: An Emerging Field in Cognitive Science. *Topics in cognitive science, 12*(1), 197–223. DOI: https://doi.org/10.1111/tops.12473

Conry-Murray, C., Waltzer, T., DeBernardi, F. C., Fossum, J. L., Haasova, S., Matthews, M. S., … Elsherif, M. M. (2023, November 17). Transparency and Validity in Coding Open-Ended Data for Quantitative Analysis. https://doi.org/10.31234/osf.io/86y9f

Damopolii, I., Lumembang, T., & İlhan, G. O. (2021). Digital comics in online learning during COVID-19: Its effect on student cognitive learning outcomes. International

Journal of Interactive Mobile Technologies (IJIM), 15(19), 33-47. DOI: https://doi.org/10.3991/ijim.v15i19.23395

Duval, S., & Tweedie, R. (2000). Trim and fill: a simple funnel‐plot–based method of testing and adjusting for publication bias in meta‐analysis. *Biometrics, 56*(2), 455-463. DOI: https://doi.org/10.1111/j.0006-341x.2000.00455.x

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *Bmj, 315*(7109), 629-634. DOI: https://doi.org/10.1136/bmj.315.7109.629

Farinella, M. (2018). Science comics' super powers. *Am Sci, 106*(4), 218. DOI: https://doi.org/10.1511/2018.106.4.218

Fleiss, J. L. (1993). Review papers: The statistical basis of meta-analysis. *Statistical methods in medical research, 2*(2), 121-145. DOI: https://doi.org/10.1177/096228029300200202

Freedman, G., Green, M. C., Kaufman, G., & Flanagan, M. (2022). Using comics and tweets to raise awareness about gender biases in STEM. *Psychology of Popular Media, 11*(1), 56. DOI: https://doi.org/10.1037/ppm0000313

Gavaler, C. (2022). The Comics Form: The Art of Sequenced Images. London, UK: Bloomsbury.

Green, M. J., & Myers, K. R. (2010). Graphic medicine: use of comics in medical education and patient care. *BMJ (Clinical research ed.), 340*, c863. DOI: https://doi.org/10.1136/bmj.c863

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends in cognitive sciences, 9*(9), 416-423. DOI: https://doi.org/10.1016/j.tics.2005.07.004

Harrer, M., Adam, S. H., Baumeister, H., Cuijpers, P., Karyotaki, E., Auerbach, R. P., ... & Ebert, D. D. (2019). Internet interventions for mental health in university students: A systematic review and meta‐analysis. *International journal of methods in psychiatric research, 28*(2), e1759. DOI: https://doi.org/10.1002/mpr.1759

IntHout, J., Ioannidis, J. P., & Borm, G. F. (2014). The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC medical research methodology*, *14*, 1-12.

Istiq'farohm N., & Mustadi, A. (2020). Improving elementary school students' creativity and writing skills through digital comics. *Ilkogretim Online, 19*(2). DOI: https://doi.org/10.17051/ilkonline.2020.689661

Kanters, S. (2022). Fixed-and random-effects models. *Meta-Research: Methods and Protocols*, 41-65.

Khoii, R., & Forouzesh, Z. (2010). Using comic strips with reading texts: Are we making a mistake. *Literacy Information and Computer Education Journal, 1*(3), 168-177.

Kim, J., Chung, M. S., Jang, H. G., & Chung, B. S. (2017). The use of educational comics in learning anatomy among multiple student groups. *Anatomical sciences education, 10*(1), 79-86. DOI: https://doi.org/10.1002/ase.1619

Knapp, G., & Hartung, J. (2003). Improved tests for a random effects meta‐regression with a single covariate. *Statistics in medicine*, *22*(17), 2693-2710. https://doi.org/10.1002/sim.1482

Kolberg, K., Larson, J., Almeida, A., Ichesco, I., Johnson, A., Van Tubbergen, M., ... & Hashikawa, A. (2021). The feasibility of using comic-based concussion discharge instructions: gauging likeability and knowledge improvement among adolescents and parents. *Pediatric emergency care, 37*(12), e1603-e1610. DOI: https://doi.org/10.1097/PEC.0000000000002133

Langan, D., Higgins, J. P. T., Jackson, D., Bowden, J., Veroniki, A. A., Kontopantelis, E., ... & Simmonds, M. (2019). A comparison of heterogeneity variance estimators in simulated random‐effects meta‐analyses. *Research Synthesis Methods, 10* (1), 83–98. https://doi.org/10.1002/jrsm.1316

Lakens, D., McLatchie, N., Isager, P. M., Scheel, A. M., & Dienes, Z. (2020). Improving inferences about null effects with Bayes factors and equivalence tests. *The Journals of Gerontology: Series B*, *75*(1), 45-57. DOI: https://doi.org/10.1093/geronb/gby065

Leiner, M., Peinado, J., Baylon, A., Lopez, I., & Pathak, I. (2018). Divide and conquer: improving parental understanding of health-related instructions using sequential pictorial instructions. *Health Education Research, 33*(2), 104-113. DOI: https://doi.org/10.1093/her/cyy004

Lestari, B., & Mustadi, A. (2020). Animated Video Media vs Comic on Storytelling Skills: Which One is More Effective?. *Journal for the Education of Gifted Young Scientists, 8*(1), 167-182. DOI: http://dx.doi.org/10.17478/jegys.664119

Lin, S. F., & Lin, H. S. (2016). Learning nanotechnology with texts and comics: the impacts on students of different achievement levels. *International Journal of Science Education, 38*(8), 1373-1391. DOI: http://dx.doi.org/10.1080/09500693.2016.1191089

Lin, S. F., Lin, H. S., Lee, L., & Yore, L. D. (2015). Are science comics a good medium for science communication? The case for public learning of nanotechnology. International *Journal of Science Education, Part B, 5*(3), 276-294. DOI: https://doi.org/10.1080/21548455.2014.941040

Lo, P., Lyu, Y. P., Chen, J. C. C., Lu, J. L., & Stark, A. J. (2022). Measuring the educational value of comic books from the school librarians' perspective: A region-wide quantitative study in Taiwan. *Journal of Librarianship and Information Science*, *54*(1), 16-33. DOI: https://doi.org/10.1177/0961000620983430

Loschky, L. C., Hutson, J. P., Smith, M. E., Smith, T. J., & Magliano, J. P. (2018). Viewing static visual narratives through the lens of the scene perception and event comprehension theory (SPECT). *In Empirical comics research* (pp. 217-238). Routledge.

Loschky, L. C., Larson, A. M., Smith, T. J., & Magliano, J. P. (2020). The scene perception & event comprehension theory (SPECT) applied to visual narratives. *Topics in cognitive science, 12*(1), 311-351. DOI: https://doi.org/10.1111/tops.12455

Lüdecke, D. (2018). ggeffects: Tidy data frames of marginal effects from regression models. *Journal of Open Source Software, 3*(26), 772. DOI: https://doi.org/10.21105/joss.00772

McCloud, S., & Martin, M. (1993). *Understanding comics: The invisible art* (Vol. 106). Northampton, MA: Kitchen sink press.

McGuinness, L. A. (2019). robvis: An R package and web application for visualising risk-of-bias assessments. *Github. URL: https://github. com/mcguinlu/robvis [accessed 2023-09-18].*

McGuinness, L. A., & Higgins, J. P. (2021). Risk‐of‐bias VISualization (robvis): an R package and Shiny web app for visualizing risk‐of‐bias assessments. *Research synthesis methods, 12*(1), 55-61.

Mayer, R. E. (2002). Multimedia learning. In *Psychology of learning and motivation* (Vol. 41, pp. 85-139). Academic Press.

Mikolajewicz, N., & Komarova, S. V. (2019). Meta-analytic methodology for basic research: a practical guide. *Frontiers in physiology, 10*, 203. DOI: https://doi.org/10.3389/fphys.2019.00203

Muzumdar, J. M., & Pal, S. (2018). Use of a Comic Book to Assist Pharmacy Student Learning of the Affordable Care Act (ACA). *Innovations in Pharmacy, 9*(2), 1. DOI: https://doi.org/10.24926/iip.v9i2.947

Newcombe, N. (2013). Seeing relationships: Using spatial thinking to teach science, mathematics, and social studies. *American Educator, 37*(1), 26.

Newcombe, N. (2017). *Harnessing Spatial Thinking to Support Stem Learning*. OECD Education Working Papers, No. 161. OECD Publishing. https://dx.doi.org/10.1787/7d5dcae6-en

O'keefe, J., & Nadel, L. (1979). Précis of O'Keefe & Nadel's The hippocampus as a cognitive map. *Behavioral and Brain Sciences*, *2*(4), 487-494. DOI: https://doi.org/10.1017/S0140525X00063949

Owens, K. A., Eno, V., Abrams, J., & Bedney, D. (2020). Comic-Con: can comics of the constitution enable meaningful learning in political science?. *PS: Political Science & Politics, 53*(1), 161-166. DOI: https://doi.org/10.1017/S1049096519001008

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *International journal of surgery, 88*, 105906. DOI: https://doi.org/10.1016/j.ijsu.2021.105906

Pagkratidou, M., Galati, A., & Avraamides, M. N. (2023, November 11). Constructing spatial situation models during reading comics vs texts. DOI: https://doi.org/10.31234/osf.io/tjdm3

Rodriguez, L., & Lin, X. (2016). The impact of comics on knowledge, attitude and behavioural intentions related to wind energy. *Journal of Visual Literacy, 35*(4), 237-252. DOI: https://doi.org/10.1080/1051144X.2016.1278090

Rücker, G., & Schwarzer, G. (2021). Beyond the forest plot: The drapery plot. *Research synthesis methods*, *12*(1), 13-19. https://doi.org/10.1002/jrsm.1410

Kelly, M., & Brier, S. (2023, January). Roy Rosenzweig Center for History and New Media Reception. In *136th Annual Meeting (January 5-8, 2023)*. AHA.

Schnotz, W., & Rasch, T. (2005). Enabling, facilitating, and inhibiting effects of animations in multimedia learning: Why reduction of cognitive load can have negative results on learning. *Educational Technology Research and Development*, *53*(3), 47-58.

Shimazaki, T., Iio, M., Uechi, H., & Takenaka, K. (2021). Emotional experiences of reading health educational manga encouraging behavioral changes: a non-randomized controlled trial. *Health Psychology and Behavioral Medicine, 9*(1), 398-421. DOI: DOI: https://doi.org/10.1080/21642850.2021.1921583

Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2020). Specification curve analysis. Nature Human Behaviour, 4 (11), 1208-1214. https://doi.org/10.1038/s41562-020-0912-z

Sivaratnam, C. S., Newman, L. K., Tonge, B. J., & Rinehart, N. J. (2015). Attachment and emotion processing in children with autism spectrum disorders: Neurobiological, neuroendocrine, and neurocognitive considerations. *Review Journal of Autism and Developmental Disorders, 2,* 222-242. DOI: https://doi.org/10.1007/s40489-015-0048-7

Spiegel, A. N., McQuillan, J., Halpin, P., Matuk, C., & Diamond, J. (2013). Engaging teenagers with science through comics. *Research in science education, 43*, 2309-2326. DOI: https://doi.org/10.1007/s11165-013-9358-x

Sousanis, N. (2015). Unflattening. Harvard University Press.

Takashi Shimazaki, Misa Iio, Hiroaki Uechi & Koji Takenaka (2021). Emotional experiences of reading health educational manga encouraging behavioral changes: a nonrandomized controlled trial, *Health Psychology and Behavioral Medicine, 9*:1, 398-421, DOI: https://doi.org/10.1080/21642850.2021.1921583

Team, R. C. (2021). R: A language and environment for statistical computing. Published online 2020. *Supplemental Information References S, 1*, 371-78.

Topkaya, Y., Batdı, V., Burak, D., & Özkaya, A. (2023). The Effectiveness of Using Comics in Education: A Meta-analytic and Meta-thematic Analysis Study. *Ahmet Keleşoğlu Eğitim Fakültesi Dergisi, 5*(3), 922-940.

Topkaya, Y., & Doğan, Y. (2020). The Effect of Educational Comics on Teaching Environmental Issues and Environmental Organizations Topics in 7th Grade Social Studies Course: A Mixed Research. *Education & Science/Egitim ve Bilim, 45*(201). DOI: https://doi.org/10.15390/EB.2019.8575

Van Dijk, T. A., & Kintsch, W. (1983). Strategies of discourse comprehension.

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2013). Three-level meta-analysis of dependent effect sizes. *Behavior research methods, 45,* 576-594. DOI: https://doi.org/10.3758/s13428-012-0261-6

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of statistical software, 36,* 1-48. DOI: https://doi.org/10.18637/jss.v036.i03

Yang, G. (2008). Graphic novels in the classroom. *Language Arts, 85*(3), 185–192

Zhu, C., Leung, C. O. Y., Lagoudaki, E., Velho, M., Segura-Caballero, N., Jolles, D., Duffy, G., Maresch, G., Pagkratidou, M., & Klapwijk, R. (2023). Fostering spatial ability development in and for authentic STEM learning. *Frontiers in Education, 8*, 1138607. https://doi.org/10.3389/feduc.2023.1138607

Zwaan, R. A., Radvansky, G. A., Hilliard, A. E., & Curiel, J. M. (1998). Constructing multidimensional situation models during reading. *Scientific studies of reading, 2*(3), 199-220. DOI: https://doi.org/10.1207/s1532799xssr0203_2

Walsh, E. I., Sargent, G. M., & Grant, W. J. (2021). Not just a pretty picture: Scientific fact visualisation styles, preferences, confidence and recall. *Information Visualization, 20*(2-3), 138-150. DOI: https://doi.org/10.1177/14738716211027587

Wang, Z., Wang, S., Farinella, M., Murray-Rust, D., Henry Riche, N., & Bach, B. (2019, May). Comparing effectiveness and engagement of data comics and infographics. *In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-12).DOI: https://doi.org/10.1145/3290605.3300483

Westerhausen, R., & Papadatou-Pastou, M. (2022). Handedness and midsagittal corpus callosum morphology: a meta-analytic evaluation. *Brain Structure and Function, 227*(2), 545-559. DOI: https://doi.org/10.1007/s00429-021-02431-4

White-Schwoch, T. (2011). Comprehending comics and graphic novels: Watchmen as a case for cognition. *SANE journal: Sequential Art Narrative in Education, 1*(2), 2.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D. A., François, R., ... & Yutani, H. (2019). Welcome to the Tidyverse. *Journal of open source software, 4*(43), 1686. DOI: https://doi.org/10.21105/joss.01686

Wiseman, R., Collver, J., Worth, R. and Watt, C. (2021). Hocus Pocus: using comics to promote skepticism about the paranormal, *JCOM, 20*(02). DOI: https://doi.org/10.22323/2.20020204

Data Supplement

A. PRISMA 2020 abstract checklist

Section and Topic	Item #	Checklist item	Reported (Yes/No)
TITLE			
Title	1	Identify the report as a systematic review.	yes
BACKGROUND			
Objectives	2	Provide an explicit statement of the main objective(s) or question(s) the review addresses.	yes
METHODS			
Eligibility criteria	3	Specify the inclusion and exclusion criteria for the review.	yes
Informatio n sources	4	Specify the information sources (e.g. databases, registers) used to identify studies and the date when each was last searched.	yes
Risk of bias	5	Specify the methods used to assess risk of bias in the included studies.	yes
Synthesis of results	6	Specify the methods used to present and synthesise results.	yes
RESULTS			

Included studies	7	Give the total number of included studies and participants and summarise relevant characteristics of studies.	N/A
Synthesis of results	8	Present results for main outcomes, preferably indicating the number of included studies and participants for each. If meta-analysis was done, report the summary estimate and confidence/credible interval. If comparing groups, indicate the direction of the effect (i.e. which group is favoured).	N/A
DISCUSSION			
Limitations of evidence	9	Provide a brief summary of the limitations of the evidence included in the review (e.g. study risk of bias, inconsistency and imprecision).	N/A
Interpretation	10	Provide a general interpretation of the results and important implications.	N/A
OTHER			
Funding	11	Specify the primary source of funding for the review.	yes
Registration	12	Provide the register name and registration number.	yes

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

B. PRISMA 2020 statement checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	In the title
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	Data Supplement A
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Manuscript: Introduction
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Manuscript: Introduction

METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Manuscript: Method
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.	Manuscript: Method
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Manuscript: Method
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.	Manuscript: Method
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.	Manuscript: Method
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.	Manuscript: Method
	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.	Manuscript: Method
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.	Manuscript: Method

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.	Manuscript: Method
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)).	Manuscript: Method
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Manuscript: Method
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Manuscript: Method
	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.	Manuscript: Method
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	Manuscript: Method
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	Manuscript: Method
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	Manuscript: Method
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Manuscript: Method

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.	N/A
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	N/A
Study characteristics	17	Cite each included study and present its characteristics.	N/A
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	N/A
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.	N/A
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	N/A
	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.	N/A
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	N/A
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	N/A

Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	N/A
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	N/A
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	N/A
	23b	Discuss any limitations of the evidence included in the review.	N/A
	23c	Discuss any limitations of the review processes used.	N/A
	23d	Discuss implications of the results for practice, policy, and future research.	N/A
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.	N/A
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	N/A
	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.	Manuscript
Competing interests	26	Declare any competing interests of review authors.	Manuscript
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.	MetaArXiv OSF

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