An umbrella review of the benefits and risks associated with youths’ interactions with electronic screens

Taren Sanders*1, Michael Noetel2, Philip Parker1, Borja Del Pozo Cruz3,14,15, Stuart Biddle4,13, Rimante Ronto5, Ryan Hulteen6, Rhiannon Parker7, George Thomas8, Katrien De Coker9, Jo Salmon10, Kylie Hesketh10, Nicole Weeks1, Hugh Arnott1, Emma Devine11, Roberta Vasconcellos1, Rebecca Pagano12, Jamie Sherson12, James Conigrave1, & Chris Lonsdale1

1 Institute for Positive Psychology and Education, Australian Catholic University, North Sydney, Australia
2 School of Psychology, University of Queensland, Brisbane, Australia
3 Department of Sport Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark
4 Centre for Health Research, University of Southern Queensland, Springfield, Australia
5 Department of Health Sciences, Faculty of Medicine, Health and Human Sciences, Macquarie University, Macquarie Park, Australia
6 School of Kinesiology, Louisiana State University, Baton Rouge, USA
7 The Centre for Social Impact, University of New South Wales, Sydney, Australia
8 The University of Queensland, Health and Wellbeing Centre for Research Innovation, School of Human Movement and Nutrition Sciences, Brisbane, Australia
9 Department of Movement and Sport Science, Ghent University, Ghent, Belgium
Author Note

Correspondence concerning this article should be addressed to Taren Sanders*, 33 Berry St, North Sydney, NSW, Australia. E-mail: Taren.Sanders@acu.edu.au
Abstract

The influence of electronic screens on children and adolescents’ health and education is not well understood. In this prospectively registered umbrella review (PROSPERO; CRD42017076051), we harmonised effects from 102 meta-analyses (2,451 primary studies; 1,937,501 participants) on screen time and outcomes. 43 effects from 32 meta-analyses met our criteria for statistical certainty. Meta-analyses of associations between screen use and outcomes showed small-to-moderate effects (range: $r = -0.14$ to 0.33). In education, results were mixed; for example, screen use was negatively associated with literacy ($r = -0.14$, 95% confidence interval [CI] -0.20 to -0.09, $p = <0.001$, $k = 38$, $N = 18,318$), but this effect was positive when parents watched with their children ($r = 0.15$, 95% CI 0.02 to 0.28, $p = 0.028$, $k = 12$, $N = 6,083$). In health, we found evidence for several small negative associations; for example, social media was associated with depression ($r = 0.12$, 95% CI 0.05 to 0.19, $p = <0.001$, $k = 12$, $N = 93,740$). Limitations include a limited number of studies for each outcome, medium-to-high risk of bias in 95/102 included meta-analyses and high heterogeneity (17/22 in education and 20/21 in health with $I^2 > 50$%). We recommend that caregivers and policymakers carefully weigh the evidence for potential harms and benefits of specific types of screen use.

Keywords:

Word count: 5312
An umbrella review of the benefits and risks associated with youths’ interactions with electronic screens

Introduction

In the 16th century, hysteria reigned around a new technology that threatened to be “confusing and harmful” to the mind. The cause of such concern? The widespread availability of books brought about by the invention of the printing press. In the early 19th century, concerns about schooling “exhausting the children’s brains” followed, with the medical community accepting that excessive study could be a cause of madness. By the 20th century, the invention of the radio was accompanied by assertions that it would distract children from their reading (which by this point was no longer considered confusing and harmful) leading to impaired learning.

Today, the same arguments that were once levelled against reading, schooling, and radio are being made about screen use (e.g., television, mobile phones, and computers). Excessive screen use is the number one concern parents in Western countries have about their children’s health and behaviour, ahead of nutrition, bullying, and physical inactivity. Yet, the evidence to support parents’ concerns is inadequate. A Lancet editorial suggested that, “Our understanding of the benefits, harms, and risks of our rapidly changing digital landscape is sorely lacking.”

While some forms of screen use (e.g., excessive television viewing) may be detrimental to health and wellbeing, evidence for other forms of screen exposure (e.g., video games or online communication, such as Zoom™) remains less certain and, in some cases, may even be beneficial. Thus, according to a Nature Human Behaviour editorial, research to determine the effect of screen exposure on youth is “a defining question of our age.” With concerns over the impact of screen use including education, health, social development, and psychological well-being, an overview that identifies potential benefits and risks is needed.
Citing the negative associations between screen use and health (e.g., increased risk of
obesity) and health-related behaviours (e.g., sleep), guidelines from the World Health
Organization\textsuperscript{12} and numerous government agencies\textsuperscript{13,14} and statements by expert groups\textsuperscript{15}
have recommended that young people’s time spent using electronic media devices for
entertainment purposes should be limited. For example, the Australian Government
guidelines regarding sedentary behaviour recommend that young children (under the age of
two) should not spend any time watching screens. They also recommend that children aged
2-5 years should spend no more than one hour engaged in recreational sedentary screen use
per day, while children aged 5-12 and adolescents should spend no more than two hours.
However, recent evidence suggests that longer exposures may not have adverse effects on
children’s behaviour or mental health—and might, in fact, benefit their well-being—as long
as exposure does not reach extreme levels (e.g., 7 hours per day)\textsuperscript{16}. Some research also
indicates that content (e.g., video games vs television programs) plays an important role in
determining the potential benefit or harm of youths’ exposure to screen-based media.\textsuperscript{17}
Indeed, educational screen use is positively related to educational outcomes.\textsuperscript{18} This evidence
has led some researchers to argue that a more nuanced approach to screen use guidelines is
required.\textsuperscript{19}

In 2016, the American Academy of Pediatrics used a narrative review to examine the
benefits and risks of children and adolescents’ electronic media\textsuperscript{20} as a basis for updating their
guidelines about screen use.\textsuperscript{15} Since then, a large number of systematic reviews and
meta-analyses have provided evidence about the potential benefits and risks of screen use.
While there have been other overviews of reviews on screen use, these have tended to focus
on a single domain (e.g., health\textsuperscript{21}), focus on a particular exposure (e.g., social media\textsuperscript{22,23}) or
provide only a narrative summary of the literature.\textsuperscript{24} Focusing on a single domain or
exposure makes it difficult to understand what trade-offs are involved in any guidelines
around screen use. For example, prohibiting screen use might reduce exposure to advertising
but may also thwart learning opportunities from interactive educational tools. Reviews on
either of these exposures or outcomes would likely miss being able to quantify these trade-offs. Overviews are one method of evidence synthesis that helps address these trade-offs, by providing ‘user-friendly’ summaries of a field of research. These overviews provide a reference point for the field and allow for easier comparison of risks and benefits for the same behaviour. By analogy, reading is a sedentary behaviour, and only by comparing the health risks against the educational benefits can researchers and policymakers make clear recommendations about what young people should do.

In order to synthesise the evidence and support further evidence-based guideline development and refinement, we reviewed published meta-analyses examining the effects of screen use on children and youth. This review synthesises evidence on any outcome of electronic media exposure. We deliberately did not pre-specify outcomes, in order to get a list of areas where there is meta-analytical evidence. Adopting this broad approach allowed us to provide a holistic perspective on the associations between screen time and different aspects of children’s lives. By synthesising across life domains (e.g., school and home), this review provides evidence to inform guidelines and advice for parents, teachers, pediatricians and other professionals in order to maximise human functioning.
Results

The searches yielded 50,649 results, of which 28,675 were duplicates. After screening titles and abstracts, we assessed 2,557 full-texts for inclusion. Of those, 217 met the inclusion criteria and we extracted the data from all of these meta-analyses. Figure 1 presents the full results of the selection process.

The most frequently reported exposures were physically active video games (n = 31), general screen use (n = 27), general TV programs and movies (n = 20), and screen-based interventions to promote health (n = 14). Supplementary File 1 provides a list of all exposures identified. The most frequently reported outcomes were body composition (n = 30), general learning (n = 24), depression (n = 13), and general literacy (n = 12). Of the 273 unique exposure/outcome combinations, 241 occurred in only one review, with 23 appearing twice, and 9 appearing three or more times. Full characteristics of the included studies are provided in Supplementary File 2. After removing reviews with duplicate exposure/outcome combinations, our process yielded 252 unique effect/outcome combinations (retaining multiple effects for different age groups or study designs) contributed from 102 reviews. These effects represent the findings of 2,451 primary studies, involving 1,937,501 participants. The characteristics of the included effects are available in Supplementary File 3.

TABLE 1

The quality of the included meta-analyses was mixed (see Table 1). Most assessed heterogeneity (n low risk = 93/102, 91% of meta-analyses), reported the characteristics of the included studies (n low risk = 86/102, 84%), and used a comprehensive and systematic search strategy (n low risk = 71/102, 70%). Most reviews did not clearly report if their eligibility criteria were predefined (n unclear = 71/102, 70%). Many papers also did not complete dual independent screening of abstracts and full text (n high risk = 20/102, 20%) or did not clearly report the method of screening (n unclear = 37/102, 36%). A similar trend
was observed for dual independent quality assessment \( n \) high risk = 52/102, 51%; \( n \) high risk = 19/102, 19%). Overall, only 7 meta-analyses were graded as low risk of bias on all criteria.

Education Outcomes

There were 88 unique effects associated with education outcomes, including general learning outcomes, literacy, numeracy, and science. We removed 28 effects that did not provide individual study-level data, 19 effects with samples < 1,000, and 19 effects with a significant Egger’s test or insufficient studies to conduct the test. Effects not meeting one or more of these standards are presented in Supplementary File 4. The remaining 22 effects met our criteria for statistical credibility and are described in Figure 2. These 22 effects came from 17 meta-analytic reviews analysing data from 337 empirical studies with 262,497 individual participants.

Among the statistically credible effects, general screen use \( (r = -0.11, 95\% \text{ confidence interval [CI]} -0.24 \text{ to } 0.01, p = 0.071, k = 18, N = 13,100) \), television viewing \( (r = -0.10, 95\% \text{ CI } -0.15 \text{ to } -0.04, p = <0.001, k = 18, N = 62,135) \), and video games \( (r = -0.08, 95\% \text{ CI } -0.12 \text{ to } -0.04, p = <0.001, k = 10, N = 4,276) \) were all negatively associated with learning. E-books that included narration \( (r = 0.11, 95\% \text{ CI } 0.05 \text{ to } 0.17, p = <0.001, k = 50, N = 2,288) \), as well as touch screen education interventions \( (r = 0.21, 95\% \text{ CI } 0.15 \text{ to } 0.28, p = <0.001, k = 79, N = 5,810) \), and augmented reality education interventions \( (r = 0.33, 95\% \text{ CI } 0.25 \text{ to } 0.42, p = <0.001, k = 15, N = 1,474) \) were positively associated with learning. General screen use was negatively associated with literacy outcomes \( (r = -0.14, 95\% \text{ CI } -0.20 \text{ to } -0.09, p = <0.001, k = 38, N = 18,318) \). However, if the screen use involved co-viewing (e.g., watching with a parent; \( r = 0.15, 95\% \text{ CI } 0.02 \text{ to } 0.28, p = 0.028, k = 12, N = 6,083) \), or the content of television programs was educational \( (r = 0.13, 95\% \text{ CI } 0.03 \text{ to } 0.23, p = 0.012, k = 13, N = 1,955) \), the association with literacy was positive and significant at the 95% confidence level (weak evidence). Numeracy outcomes were positively associated with screen-based mathematics interventions \( (r = 0.27, 95\% \text{ CI } 0.21 \text{ to } 0.33, p = \)
<0.001, k = 85, N = 36,793) and video games that contained numeracy content (r = 0.32, 95% CI 0.21 to 0.43, p = <0.001, k = 25, N = 2,008).

As shown in Figure 2, most of the credible results (13 of 22 effects) showed statistically significant associations, with 99.9% confidence intervals not encompassing zero (strong evidence). The remaining six associations were significant at the 95% confidence level (weak evidence). All credible effects related to education outcomes were small-to-moderate.

Screen-based interventions designed to influence an outcome (e.g., a computer based program designed to enhance learning; r = 0.21, 95% CI 0.15 to 0.28, p = <0.001, k = 79, N = 5,810) tended to have larger effect sizes than exposures that were not specifically intended to influence any of the measured outcomes (e.g., the association between television viewing and learning; r = -0.10, 95% CI -0.15 to -0.04, p = <0.001, k = 18, N = 62,135). The largest effect size observed was for augmented reality-based education interventions on general learning (r = 0.33, 95% CI 0.25 to 0.42, p = <0.001, k = 15, N = 1,474). Most effects showed high levels of heterogeneity (17 of 22 with $I^2 > 50$).

Health-related Outcomes

We identified 163 unique outcome-exposure combinations associated with health or health-related behaviour outcomes. We removed 39 effects that did not provide individual study-level data, 50 effects with samples < 1,000, and 53 effects with a significant Egger’s test or insufficient studies to conduct the test. No remaining studies had statistically significant tests for excess significance. Effects not meeting one or more of these standards are presented in Supplementary File 5. The remaining 21 meta-analytic associations met our criteria for credible evidence and are described below (see also Figure 3). These 21 effects came from 15 meta-analytic reviews analysing data from 344 empirical studies with 859,562 individual participants.

Digital advertising of unhealthy foods—both traditional advertising (r = 0.23, 95% CI
0.10 to 0.37, \( p = <0.001, k = 13, N = 1,756 \) and video games developed by a brand for promotion \( (r = 0.18, 95\% \text{ CI} 0.10 \text{ to } 0.25, p = <0.001, k = 15, N = 3,842) \)—were associated with higher unhealthy food intake. Social media use and sexual content were positively associated with risky behaviors \( (e.g., \text{social media and risky sexual behaviour}; r = 0.21, 95\% \text{ CI} 0.14 \text{ to } 0.28, p = <0.001, k = 14, N = 23,096) \). Television viewing was negatively correlated with sleep duration, but with stronger evidence only observed for adolescents \( (r = -0.06, 95\% \text{ CI} -0.10 \text{ to } -0.01, p = 0.018, k = 10, N = 9,798) \). Both television and video games were associated with body composition \( (e.g., \text{television} r = 0.06, 95\% \text{ CI} 0.03 \text{ to } 0.10, p = <0.001, k = 12, N = 3,196) \). Screen-based interventions which target health behaviours appeared mostly effective.

Across the health outcomes, most \( (14 \text{ of } 21) \) effects were statistically significant at the 99.9\% confidence interval level, with the remaining four significant at 95\% confidence. However, most of the credible effects exhibited high levels of heterogeneity, with all but two having \( I^2 > 75\% \). Additionally, most effects were small, with the association between internet use and depression the largest at \( r = 0.25 \ (95\% \text{ CI} 0.22 \text{ to } 0.27, p = <0.001, k = 118, N = 527,696) \). Most of the effect sizes \( (17/21) \) had an absolute value of \( r < 0.2 \).

**Discussion**

The primary goal of this review was to provide a holistic perspective on the association between screen use and a broad range of health- and education-related aspects of children’s lives. We found that when meta-analyses examined general screen use, and did not specify the content, context or device, there was strong evidence showing potentially harmful associations with general learning, literacy, body composition, and depression. However, when meta-analyses included a more nuanced examination of exposures, a more complex picture appeared.

As an example, consider children watching television programs—an often cited form of
screen use harm. We found evidence for a small association with poorer academic
performance and literacy skills for general television watching\textsuperscript{29}. However, we also found
evidence that if the content of the program was educational, or the child was watching the
program with a parent (i.e., co-viewing), this exposure was instead associated with better
literacy\textsuperscript{143}. Thus, parents may play an important role in selecting content that is likely to
benefit their children or, perhaps, interact with their children in ways that may foster
literacy (e.g., asking their children questions about the program). Similar nuanced findings
were observed for video games. The credible evidence we identified showed that video game
playing was associated with poorer body composition and learning\textsuperscript{29,173}. However, when the
video game were designed specifically to teach numeracy, playing these games showed
learning benefits\textsuperscript{52}. One might expect that video games designed to be physically active
could confer health benefits, but none of the meta-analyses examining this hypothesis met
our thresholds for statistical credibility (see Supplementary Files 4 & 5) therefore this
hypothesis could not be addressed.

Social media was one type of exposure that showed consistent—albeit
small—associations with poor health, with no indication of potential benefit. Social media
showed strong evidence of harmful associations with risk taking in general, as well as unsafe
sex and substance abuse\textsuperscript{218}. These results align with meta-analytic evidence from adults
indicating that social media use is also associated with increased risk of depression\textsuperscript{214,233}.
Recent evidence from social media companies themselves suggest there may also be negative
effects of social media on the mental health of young people, especially teenage girls\textsuperscript{243}.

One category of exposure appeared to be consistently associated with benefits:
screen-based interventions designed to promote learning or health behaviours. This finding
indicates that interventions can be effectively delivered using electronic media platforms, but
does not necessarily indicate that screens are more effective than other methods (e.g.,
face-to-face, printed material). Rather, it reinforces that the content of the screen use may
be the most important aspect. The way that a young person interacts with digital screens may also be important. We found evidence that touch screens had strong evidence for benefits on learning\textsuperscript{230} as did augmented reality\textsuperscript{207}.

Largely owing to a small number of studies or missing individual study data, there were few age-based conclusions that could be drawn from reviews which met our criteria for statistical certainty. Given the differences in development across childhood and adolescence and the different ways children of various ages use screens, further examination of age-based differences is needed. However, in the absence of this work, our study has shown how children are affected by screens in general.

Among studies that met our criteria for statistical certainty heterogeneity was high, with almost all effects having $I^2 > 50\%$. Much of this heterogeneity is likely explained by differences in measures across pooled studies, or in some cases, the generic nature of some of the exposures. For example, “TV programs and movies” covers a substantial range of content, which may explain the heterogeneous association with education outcomes.

Our results have several implications for policy and practice. Broadly, our findings align with the recommendations of others who suggest that current guidelines may be too simplistic, mischaracterise the strength of the evidence, or do not acknowledge the important nuances of the issue\textsuperscript{244–246}. Our findings suggest that screen use is a complex issue, with associations based not just on duration and device type, but also on the content and the environment in which the exposure occurs. Many current guidelines simplify this complex relationship as something that should be minimised.\textsuperscript{12,13} We suggest that future guidelines need to embrace the complexity of the issue, to give parents and clinicians specific information to weigh the pros and cons of interactions with screens.

Given our results, we support the continuing trend of guidelines moving away from recommendations to reduce ‘screen use’, and instead focusing on the type of screen use. For
example, we suggest that guidelines should discourage high levels of social media and internet use. Guidelines may also consider adapting recommendations that promote the use of educational apps and video games, although these recommendations need to be balanced against the (very small) risks to adiposity.\textsuperscript{151}

Our results also have implications for future research. Screen use research is extensive, varied, and rapidly growing. Reviews tended to be general (e.g., all screen use) and even when more targeted (e.g., social media) nuances related to specific content (e.g., Instagram vs Facebook) have not been meta-analysed or have not produced credible evidence. Fewer than 20\% of the effects identified met our criteria for statistical credibility. Most studies which did not meet our criteria failed to provide study-level data (or did not provide sufficient data, such as including effect estimates but not sample sizes). Newer reviews were more likely to provide this information than older reviews, but it highlights the importance of data and code sharing as recommended in the PRISMA guidelines.\textsuperscript{247} When study level data was available, many effects were removed because the pooled sample size was small, or because there were fewer than ten studies on which to perform an Egger’s test. It seems that much of the current screen use research is small in scale, and there is a need for larger, high-quality studies.

Our results highlight the need for the field to more carefully consider if the term ‘screen use’ remains appropriate for providing advice to parents. Instead, our results suggest that more nuanced and detailed descriptions of the behaviours to be modified may be required. Rather than suggesting parents limit ‘screen use’, for example, it may be better to suggest that parents promote interactive educational experiences but limit exposure to advertising.

Screen use research has a well-established measurement problem, which impacts the individual studies of this umbrella review. The vast majority of screen use research relies on self-reported data, which not only lacks the nuance required for understanding the effects of screen use, but may also be inaccurate. In one systematic review on screen use and sleep,\textsuperscript{7} 66
of the 67 included studies used self-reported data for both the exposure and outcome variable. It has been established that self-reported screen use data has questionable validity. In a meta-analysis of 47 studies comparing self-reported media use with logged measures, Parry et al. found that the measures were only moderately correlated \( r = 0.38 \), with self-reported problematic usage fairing worse \( r = 0.25 \). Indeed, of 622 studies which measured the screen use of 0—6 year-olds, only 69 provided any sort of psychometric properties for their measure, with only 19 studies reporting validity. While some researchers have started using newer methods of capturing screen behaviours—such as wearable cameras or device-based loggers—these are still not widely adopted. It may be that the field of screen use research cannot be sufficiently advanced until accurate, validated, and nuanced measures are more widely available and adopted.

There were a number of strengths and limitations to our work. Our primary goal for this umbrella review was to provide a high-level synthesis of screen use research, by examining a range of exposures and the associations with a broad scope of outcomes. Our results represent the findings from 2,451 primary studies comprised of 1,937,501 participants. To ensure findings could be compared on a common metric, we extracted and reanalysed individual study data where possible.

Our high-level approach limits the feasibility of examining fine-grained details of the individual studies. For example, we did not examine moderators beyond age, nor did we rate the risk of bias for the individual studies. Thus, our assessment of evidence quality was restricted to statistical credibility, rather than a more complete assessment of quality (e.g., GRADE). As such, we made decisions regarding the credibility of evidence, where others may have used different thresholds or metrics. In addition, when faced with duplicate outcome/exposure combinations we chose to keep the one with the largest pooled sample size, assuming that this would capture the most comprehensive and most recent review. Inspection of the excluded effect sizes suggests that this decision was not that impactful: our
results would have been almost exactly the same has we used the number of included studies
\( (k) \) or the most recent review by publication year. However, we provide the complete results
in Supplementary Files 4 & 5, along with the dataset (Supplementary File 6) for others to
consider alternative criteria.

Our high-level approach also means that we could not engage with the specific
mechanisms behind each association, and as such, we cannot make claims on the directions
of causality. These likely depend on the specific exposure and outcome. It is tempting to
draw inferences that the associations are due to screen use causing these outcomes, but we
cannot rule out reverse causality, a third variable, or some combination of influences. Many
of the individual reviews go into more detail about the strength of the evidence for causal
associations, but those judgements were difficult to synthesise across more than 200 reviews.
Readers who wish to more deeply understand one specific relationship are directed to the
cited review for that effect, where the authors could engage more deeply with the
mechanisms.

We converted all effect sizes to a common metric (Pearson’s r) to allow for comparisons
of magnitude, but acknowledge that this assumes a linear relationship between the variables.
Some previous research suggests that associations are typically linear.\(^1^8\) However, others
have identified instances where non-linear relationships exist, especially for very high levels
of screen use.\(^1^7,2^5^3,2^5^4\) Additionally, our conversion may not always adequately account for
differences in study design or measures of exposures and outcomes. Care is needed, therefore,
when interpreting the effect sizes. In addition, reviews provide only historical evidence which
may not keep up with the changing ways children can engage with screens. While our
synthesis of the existing evidence provides information about how screens might have
influenced children in the past, it is difficult to know if these findings will translate to new
forms of technology in the future.

Screen use is a topic of significant interest, as shown by the wide variety of academic
domains involved, parents’ concerns, and the growing pervasiveness into society. Our findings showed that screen use is associated with both positive (e.g., educational video games were associated with improved literacy) and negative (e.g., general screen use was associated with poorer body composition) outcomes. Based on our findings, we recommend that parents, teachers, and other caregivers need to carefully weigh the evidence for pros and cons of each specific activity for potential harms and benefits. However, our findings also lead us to suggest that in order to aid caregivers to make this judgement, researchers need to conduct more careful and nuanced measurement and analysis of screen use, with less emphasis on measures that aggregate screen use and instead focus on the content, context, and environment in which the exposure occurs.

Methods

We prospectively registered our methods on the International Prospective Register of Systematic Reviews (PROSPERO; CRD42017076051) in October 2017. We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.247

Eligibility criteria. Population: To be eligible for inclusion, meta-analyses needed to include meta-analytic effect sizes for children or adolescents (age 0-18 years). We included meta-analyses containing studies that combined data from adults and youth if meta-analytic effect size estimates specific to participants aged 18 years or less could be extracted (i.e., the highest mean age for any individual study included in the meta-analysis was < 18 years). A meta-analysis was still included if the age range exceed 18 years, provided that the mean age was less than 18. We excluded meta-analyses that only contained evidence gathered from adults (age >18 years).

Exposure: We included meta-analyses examining all types of electronic screens including (but not necessarily limited to) television, gaming consoles, computers, tablets, and mobile phones. We also included analyses of all types of content on these devices,
including (but not necessarily limited to) recreational content (e.g., television programs, movies, games), homework, and communication (e.g., video chat). In this review we focused on electronic media exposure that would be considered typical for children and youth. That is, exposure that may occur in the home setting, or during schooling. Consistent with this approach, we excluded technology-based treatments for clinical conditions. However, we included studies examining the effect of screen exposure on non-clinical outcomes (e.g., learning) for children and youth with a clinical condition. For example, a meta-analysis of the effect of television watching on learning among adolescents diagnosed with depression would be included. However, a meta-analysis of interventions designed to treat clinical depression delivered by a mobile phone app would be excluded.

Outcomes: We included all reported outcomes on benefits and risks.

Publications: We included meta-analyses (or meta-regressions) of quantitative evidence. To be included, meta-analyses needed to analyse data from studies identified in a systematic review. For our purposes, a systematic review was one in which the authors attempted to acquire all the research evidence that pertained to their research question(s). We excluded meta-analyses that did not attempt to summarise all the available evidence (e.g., a meta-analysis of all studies from one laboratory). We included meta-analyses regardless of the study designs included in the review (e.g., laboratory-based experimental studies, randomised controlled trials, non-randomised controlled trials, longitudinal, cross-sectional, case studies), as long as the studies in the review collected quantitative evidence. We excluded systematic reviews of qualitative evidence. We did not formulate inclusion/exclusion criteria related to the risk of bias of the review. We did, however, employ a risk of bias tool to help interpret the results. We included full-text, peer-reviewed meta-analyses published or ‘in-press’ in English. We excluded conference abstracts and meta-analyses that were unpublished.
**Information sources.** We searched records contained in the following databases: Pubmed, MEDLINE, CINAHL, PsycINFO, SPORTDiscus, Education Source, Embase, Cochrane Library, Scopus, Web of Science, ProQuest Social Science Premium Collection, and ERIC. We conducted an initial search on August 17, 2018 and refreshed the search on September 27, 2022. We searched reference lists of included papers in order to identify additional eligible meta-analyses. We also searched PROSPERO to identify relevant protocols and contacted authors to determine if these reviews have been completed and published.

**Search strategy.** The search strategy associated with each of the 12 databases can be found in Supplementary File 7. We hand searched reference lists from any relevant umbrella reviews to identify systematic meta-analyses that our search may have missed.

**Selection process.** Using Covidence software (Veritas Health Innovation, Melbourne, Australia), two researchers independently screened all titles and abstracts. Two researchers then independently reviewed full-text articles. We resolved disagreements at each stage of the process by consensus, with a third researcher employed, when needed.

**Data items.** From each included meta-analysis, two researchers independently extracted data into a custom-designed database. We extracted the following items: First author, year of publication, study design restrictions (e.g., cross-sectional, observational, experimental), region restrictions (e.g., specific countries), earliest and latest study publication dates, sample age (mean), lowest and highest mean age reported, outcomes reported, and exposures reported.

**Study risk of bias assessment.** For each meta-analysis, two researchers independently completed the National Health, Lung and Blood Institute’s Quality Assessment of Systematic Reviews and Meta-Analyses tool\(^\text{255}\) (see Table 1). We resolved disagreements by consensus, with a third researcher employed when needed. We did not assess risk of bias in the individual studies that were included in each meta-analysis.
**Effect measures.** Two researchers independently extracted all quantitative meta-analytic effect sizes, including moderation results. We excluded effect sizes which were reported as relative risk ratios or odds ratios, as meta-analyses did not contain sufficient information to meaningfully convert to a correlation. We also excluded effect size estimates when the authors did not provide a sample size. Where possible, we also extracted effect sizes from the primary studies included in each meta-analysis.

To facilitate comparisons, we converted effect sizes to Pearson’s $r$ using established formulae.\textsuperscript{256,257} Effect sizes on the original metric are provided in Supplementary File 6. Throughout the results section we interpret the size of the effects using Funder and Ozer’s guidelines:\textsuperscript{258} very small ($0.05 < r <= 0.1$), small ($0.1 < r <= 0.2$), medium ($0.2 < r <= 0.2$), large ($0.3 < r <= 0.4$), and very large ($r >= 0.4$). These are similar to other interpretations based on empirical data.\textsuperscript{259}

**Synthesis methods.** After extracting data, we examined the combinations of exposure and outcomes and removed any effects that appeared multiple times (i.e., in multiple meta-analyses, or with multiple sub-groups in the same meta-analysis), keeping the effect with the largest total sample size. In instances where effect sizes from the same combination of exposure and outcome were drawn from different age-groups (e.g., children vs adolescents), or were drawn using different study designs (e.g., cross-sectional vs longitudinal) we retained both estimates in our dataset.

We descriptively present the remaining meta-analytic effect sizes. To remove the differences in approach to meta-analyses across the reviews, we reran the effect size estimate using a random effects meta-analysis via the metafor package\textsuperscript{260} in R\textsuperscript{261} (version 4.3.0) when the meta-analysis’s authors provided primary study data associated with these effects. When required, we imputed missing sample sizes using mean imputation from the other studies within that review. From our reanalysis we also extracted $I^2$ values. To test for publication bias, we conducted Egger’s test\textsuperscript{262} when the number of studies within the review was ten or
more, and conducted a test of excess significance. We contacted authors who did not provide primary study data in their published article. Where authors did not provide data in a format that could be re-analysed, we used the published results of their original meta-analysis.

Evidence assessment criteria. Statistical Credibility: We employed a statistical classification approach to grade the credibility of the effect sizes in the literature. To be considered ‘credible’ an effect needed to be derived from a combined sample of >1,000 participants and have non-significant tests of publication bias (i.e., Egger’s test and excess significance test). We performed these analyses, and therefore the review needed to provide usable study-level data in order to be included.

Consistency of Effect within the Population: We also examined the consistency of the effect size using the $I^2$ measure. We considered $I^2 < 50\%$ to indicate effects that were relatively consistent across the population of interest. $I^2$ values of $> 50\%$ were taken to indicate an effect was potentially heterogeneous within the population.

Direction of Effect: Finally, we examined the extent to which significance testing suggested screen exposure was associated with benefit, harm, or no effect on outcomes. We used thresholds of $P < .05$ for weak evidence (i.e., 95% confidence intervals did not cross zero) and $P < 10^{-3}$ (i.e., 99.9% confidence intervals did not cross zero) for strong evidence. An effect with statistical credibility but with $P > .05$ (i.e., 95% confidence intervals included zero) was taken to indicate no association of interest.

Deviations from protocol. As described above, we have summarised the meta-analytic findings from all included systematic reviews. In our protocol, we originally planned to also conduct a narrative synthesis of all systematic reviews, even those without meta-analyses. However, we determined that combining results from the meta-analyses alone allow readers to compare relative strength of associations more easily. Readers interested in the relevant systematic reviews (i.e., without meta-analysis) can consult the list of references.
We altered our evidence assessment plan when we identified that, as written, it could not classify precise evidence of null effects (i.e., from large reviews with low heterogeneity and low risk of publication bias) as ‘credible’ because a highly-significant $P$-value was a criteria. This would have significantly harmed knowledge gained from our review as it would have restricted our ability to show where the empirical evidence strongly indicated that there was no association between screen use and a given outcome.

**Data availability statement**

All data for this review are available from the authors’ GitHub repository (https://github.com/motivation-and-Behaviour/screen_umbrella) or from the Open Science Foundation (https://osf.io/3ubqp/).

**Code availability statement**

All code used in these analyses are available on the authors’ GitHub repository (https://github.com/motivation-and-Behaviour/screen_umbrella).

**Acknowledgements**

The authors received no specific funding for this work.

**Author contributions**

TS, MN, PP, and CL conceptualised the review and drafted the manuscript. TS, MN, and PP conducted the analyses. All authors contributed to data extraction, interpretation, and editing of the manuscript.

**Competing interests**

The authors declare no conflicts of interest.
Tables

Table 1: Review characteristics and quality assessment for meta-analyses providing unique effects
Figure legends

Figure 1: PRISMA flow diagram.

Figure 2: Education outcomes. Forest plot for 22 unique effect sizes related to educational outcomes which met the criteria for statistical certainty. Findings are presented as correlations (two-sided) with both 95% and 99.9% confidence intervals.

Figure 3: Health and health-related behaviour outcomes. Forest plot for 21 unique effect sizes related to health and health-related behaviour outcomes which met the criteria for statistical certainty. Findings are presented as correlations (two-sided) with both 95% and 99.9% confidence intervals.
References


2. Bell, A. N. *The sanitarian*. vol. 11 (AN Bell, 1883).


Performance Among Children and Adolescents: A Systematic Review and Meta-analysis.  


61. Chauhan, S. A meta-analysis of the impact of technology on learning effectiveness

673

674 62. Chen, L., Ho, S. S. & Lwin, M. O. A meta-analysis of factors predicting
675 cyberbullying perpetration and victimization: From the social cognitive and media effects

677

678 63. Chen, C.-H., Shih, C.-C. & Law, V. The Effects of Competition in Digital
679 Game-Based Learning (DGBL): A Meta-Analysis. *Educational Technology Research and

681

682 64. Cheung, A. C. K. & Slavin, R. E. How features of educational technology
683 applications affect student reading outcomes: A meta-analysis. *Educational Research Review

685

686 65. Cheung, A. C. K. & Slavin, R. E. The effectiveness of educational technology
687 applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis.

689

690 66. Cheung, A. C. K. & Slavin, R. E. Effects of Educational Technology Applications
691 on Reading Outcomes for Struggling Readers: A Best-Evidence Synthesis. *Reading Research

693


697

698 68. Chodura, S., Kuhn, J.-T. & Holling, H. Interventions for Children With

699

700 69. Claussen, A. H. *et al.* All in the Family? A Systematic Review and Meta-analysis
701 of Parenting and Family Environment as Risk Factors for Attention-Deficit/Hyperactivity


102. Garzón, J. & Acevedo, J. Meta-analysis of the impact of Augmented Reality on


110. He, Z. *et al.* Effects of Smartphone-Based Interventions on Physical Activity in Children and Adolescents: Systematic Review and Meta-analysis. *JMIR mHealth and


159. Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W. & Davis, T. J. Effectiveness of virtual reality-based instruction on students’ learning outcomes in K-12 and


183. Scherer, R., Siddiq, F. & Sánchez Viveros, B. A meta-analysis of teaching and


191. Slavin, R. E. Reading Effects of IBM’s "Writing to Read" Program: A Review of


231. Yang, Q., Liu, J. & Rui, J. Association between social network sites use and mental illness: A meta-analysis. *Cyberpsychology: Journal of Psychosocial Research on*


### Table 1

**Quality assessment for studies providing unique effects**

<table>
<thead>
<tr>
<th>First Author</th>
<th>Year</th>
<th>Elig.</th>
<th>Lit.</th>
<th>Dual Screen</th>
<th>Dual Qual.</th>
<th>Studies Listed</th>
<th>Pub Bias</th>
<th>Hetero.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abrami</td>
<td>2020</td>
<td>U</td>
<td>U</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Adelantado-Renau</td>
<td>2019</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Andrade</td>
<td>2019</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Arztmann</td>
<td>2022</td>
<td>U</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Aspiranti</td>
<td>2020</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Bartel</td>
<td>2015</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>Beck Silva</td>
<td>2022</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Benavides-Varela</td>
<td>2020</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Blok</td>
<td>2002</td>
<td>U</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Bossen</td>
<td>2020</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Boyland</td>
<td>2016</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Byun</td>
<td>2018</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>Cao</td>
<td>2020</td>
<td>U</td>
<td>H</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Champion</td>
<td>2019</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Chan</td>
<td>2014</td>
<td>U</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Chauhan</td>
<td>2017</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Chen</td>
<td>2020</td>
<td>U</td>
<td>H</td>
<td>U</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Cheung</td>
<td>2012</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Cheung</td>
<td>2013</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Cho</td>
<td>2018</td>
<td>U</td>
<td>H</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
</tbody>
</table>
Table 1

Quality assessment for studies providing unique effects (continued)

<table>
<thead>
<tr>
<th>First Author</th>
<th>Year</th>
<th>Elig.</th>
<th>Lit.</th>
<th>Dual Crit. ^1</th>
<th>Search ^2</th>
<th>Dual Screen ^3</th>
<th>Dual Qual. ^4</th>
<th>Studies Listed ^5</th>
<th>Pub. Bias ^6</th>
<th>Hetero. ^7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Claussen</td>
<td>2022</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Clinton</td>
<td>2019</td>
<td>U</td>
<td>H</td>
<td>U</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Comeras-Chueca</td>
<td>2021</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Comeras-Chueca</td>
<td>2021</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Coyne</td>
<td>2018</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Cunningham</td>
<td>2021</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Cushing</td>
<td>2010</td>
<td>U</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Darling</td>
<td>2017</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Eirich</td>
<td>2022</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Feng</td>
<td>2021</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Ferguson</td>
<td>2017</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Ferguson</td>
<td>2020</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Folkvord</td>
<td>2018</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Furenes</td>
<td>2021</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Gardella</td>
<td>2017</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Garzón</td>
<td>2019</td>
<td>U</td>
<td>H</td>
<td>U</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Graham</td>
<td>2015</td>
<td>U</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Hammersley</td>
<td>2016</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Hao</td>
<td>2021</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Hassan-Saleh</td>
<td>2019</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>U</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>He</td>
<td>2021</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
</tbody>
</table>
Table 1

*Quality assessment for studies providing unique effects (continued)*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Hernandez-Jimenez</td>
<td>2019</td>
<td>U</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Hurwitz</td>
<td>2018</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Ivie</td>
<td>2020</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Janssen</td>
<td>2020</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Kates</td>
<td>2018</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Kim</td>
<td>2021</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Kroesbergen</td>
<td>2003</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Kucukalkan</td>
<td>2019</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Li</td>
<td>2010</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Li</td>
<td>2022</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Li</td>
<td>2022</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Liao</td>
<td>2008</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>Liao</td>
<td>2014</td>
<td>U</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Liu</td>
<td>2019</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Liu</td>
<td>2022</td>
<td>U</td>
<td>H</td>
<td>U</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Lu</td>
<td>2021</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Madigan</td>
<td>2020</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Major</td>
<td>2021</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Mallawaarachchi</td>
<td>2022</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Mares</td>
<td>2005</td>
<td>U</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>Mares</td>
<td>2013</td>
<td>U</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
</tr>
</tbody>
</table>
Table 1

*Quality assessment for studies providing unique effects (continued)*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Marker</td>
<td>2022</td>
<td>U</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Marshall</td>
<td>2004</td>
<td>U</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Martins</td>
<td>2019</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Martins</td>
<td>2022</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Mazeas</td>
<td>2022</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>McArthur</td>
<td>2012</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>McArthur</td>
<td>2018</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Mei</td>
<td>2018</td>
<td>U</td>
<td>H</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Merchant</td>
<td>2014</td>
<td>U</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Neitzel</td>
<td>2022</td>
<td>U</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>Oldrati</td>
<td>2020</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Paik</td>
<td>1994</td>
<td>U</td>
<td>H</td>
<td>U</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>H</td>
</tr>
<tr>
<td>Pearce</td>
<td>2016</td>
<td>U</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Peng</td>
<td>2011</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>U</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Powers</td>
<td>2013</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Prescott</td>
<td>2018</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Reynard</td>
<td>2022</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Rodriguez-Rocha</td>
<td>2019</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Sadeghirad</td>
<td>2016</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Scherer</td>
<td>2020</td>
<td>U</td>
<td>H</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Schroeder</td>
<td>2013</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
</tbody>
</table>
Table 1

*Quality assessment for studies providing unique effects (continued)*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Scionti</td>
<td>2019</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Shin</td>
<td>2019</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Shin</td>
<td>2022</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Slavin</td>
<td>2014</td>
<td>U</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>Strouse</td>
<td>2021</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Takacs</td>
<td>2014</td>
<td>H</td>
<td>L</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Takacs</td>
<td>2019</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Tekedere</td>
<td>2016</td>
<td>U</td>
<td>H</td>
<td>U</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Tokac</td>
<td>2019</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Vahedi</td>
<td>2018</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>van Ekris</td>
<td>2016</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Vannucci</td>
<td>2020</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Williams</td>
<td>1982</td>
<td>U</td>
<td>U</td>
<td>H</td>
<td>U</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>Wouters</td>
<td>2013</td>
<td>U</td>
<td>H</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Xie</td>
<td>2018</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Yin</td>
<td>2019</td>
<td>U</td>
<td>H</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Zhou</td>
<td>2020</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Note:</td>
<td>Items are from the National Health, Lung and Blood Institute’s Quality Assessment of Systematic Reviews and Meta-Analyses tool. Note that we excluded the first item of the tool. U = Unclear; L = Low; H = High</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Eligibility criteria predefined and specified</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Literature search strategy comprehensive and systematic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Dual independent screening and review</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Dual independent quality assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Included studies listed with important characteristics and results of each</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Publication bias assessed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Heterogeneity assessed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Zucker</th>
<th>2009</th>
<th>L</th>
<th>L</th>
<th>U</th>
<th>H</th>
<th>L</th>
<th>H</th>
<th>L</th>
</tr>
</thead>
</table>