DOI: 10.1111/cdev.13927

CHILD DEVELOPMENT

REVIEW

Screen media exposure and young children's vocabulary learning and development: A meta-analysis

Mengguo Jing[1] | Ting Ye[2] | Heather L. Kirkorian[3] | Marie-Louise Mares[4]

[1]Department of Psychology and Neuroscience, Boston College, Chestnut Hill, Massachusetts, USA
[2]Biostatistics, University of Washington, Seattle, Washington, USA
[3]Human Development and Family Studies, University of Wisconsin-Madison, Madison, Wisconsin, USA
[4]Communication Arts, University of Wisconsin-Madison, Madison, Wisconsin, USA

Correspondence
Mengguo Jing, Department of Psychology and Neuroscience, Boston College, 140 Commonwealth Avenue, Chestnut Hill, MA 02467, USA.
Email: jingmengguo@gmail.com

Funding information
University of Wisconsin-Madison

Abstract
This meta-analysis synthesizes research on media use in early childhood (0–6 years), word-learning, and vocabulary size. Multi-level analyses included 266 effect sizes from 63 studies (N_{total} = 11,413) published between 1988–2022. Among samples with information about race/ethnicity (51%) and sex/gender (73%), most were majority White/Non-Hispanic and between 40%–60% female. Analyses revealed a small overall positive relation between screen media exposure and vocabulary (r = .23). Experimental studies yielded a small-to-medium effect (r = .30), with stronger effects for e-books than TV/video or games/apps, and non-significant effects for video chat. In correlational studies, there was no overall association between vocabulary size and naturalistic media exposure (r = .07), with the exception of naturalistic exposure to educational media (r = .17).

Screen media have often been blamed for capturing young viewers' passive engagement and displacing time that might otherwise have been spent on more valuable activities. Such concerns are partially fueled by research showing negative associations between screen time and child outcomes, such as language, attention, and self-regulation (for a review, see Anderson & Kirkorian, 2015). At the same time, research also suggests that, given their prevalence and accessibility, screen media have the potential to break the boundary between formal and informal learning, and to help mitigate the effects of other constraints on early learning, such as economic disparities and physical distance (e.g., Cheng & Wilkinson, 2020; Verhallen & Bus, 2010). Indeed, nationwide polls revealed that U.S. parents hold mostly optimistic beliefs that their very young children can learn from high quality, age-appropriate screen media (Rideout & Robb, 2020; Vandewater et al., 2007).

Given mixed findings in the literature, the current meta-analysis examines the relation between young children's screen media use and their vocabulary learning and development, investigating the conditions that give rise to positive or negative associations. Vocabulary learning is of interest because it is a particularly strong predictor of child outcomes, including reading skills (Dickinson & Porche, 2011; Morgan et al., 2015), emergent literacy and school readiness (Niessen et al., 2011; Pace et al., 2019) as well as future literacy and academic achievement (Paris, 2005; Snow et al., 1998). Indeed, the National Reading Panel Report (2000) proposed screen media use as one possible avenue for enhancing children's vocabulary knowledge. Given that the last decade has brought an upsurge of interactive and educational-labeled media marketed as bolstering early learning (Apple, 2017; Rideout, 2017), it is critical to evaluate how screen media exposure may affect early vocabulary.

Abbreviation: SES, socioeconomic status.

Vocabulary and screen media

"Screen media" is a broad term that encompasses many different types of technology that can provide access to many different types of content. The defining element of screen media use or exposure is that the user has mediated (rather than direct) contact with events presented on some form of electronic screen with a visual component. Within this broad definition, research on children's use of screen media has considered a range of devices and experiences including television, film, and streaming video content; electronic books (e.g., with some audio and/or animated presentation of the text); digital games played on computers, touch-screen devices, or other technology; and real-time video chat interactions (e.g., via Zoom or Skype).

The causal impact of screen media use on vocabulary development has been theorized in various ways. All of them rest on research suggesting that children acquire vocabulary by observation, interaction, and contingent feedback, with most growth occurring during the preschool years in oral vocabulary (Farkas & Beron, 2004; Hirsh-Pasek et al., 2015). Based on this understanding, the core argument for negative effects is that screen use may displace or disrupt time spent in those social activities that would be more beneficial for vocabulary development (see Anderson & Hanson, 2017 for a review). A related argument is that young children are less likely to learn new words from video lessons than from in-person interactions (Krcmar et al., 2007; Roseberry et al., 2009), even when there is a live social partner on the screen in real time (Troseth et al., 2018). As such, the opportunity costs of screen time are potentially considerable during a developmental period that is particularly important for vocabulary growth.

Conversely, the core argument for positive effects is that screen use potentially provides many opportunities for exposure to language-enriching content, including content that is explicitly designed by educational experts to foster vocabulary development. A related argument is that some forms of screen use (e.g., video chat) provide many, if not all, of the conditions (e.g., social contingency) considered important for vocabulary acquisition (Roseberry et al., 2014). Other types of screen use (e.g., shared reading of e-books) can provide opportunities for parent–child interactions that enhance vocabulary acquisition (Mendelsohn et al., 2010). Thus, at least some forms of screen use and some types of content may enrich rather than disrupt vocabulary development.

Some of these issues were considered in a systematic review and meta-analysis by Madigan et al. (2020), examining non-experimental associations between screen time and language skills among children aged 12 or younger. Based on a sample of 42 studies (up to early 2019), they reported a small negative correlation ($r = -.14$) between language skills and hours of screen time, regardless of whether that screen time only included television viewing or also included use of other devices such as computers, mobile phones, and video games. However, language skills were weakly positively correlated with exposure to educational content ($r = .13$) and time spent co-viewing video content with a parent or caregiver ($r = .16$), with the latter co-viewing association stronger for samples with more boys.

The current meta-analysis builds on this work in a number of ways, adding recent studies (i.e., 9 studies available between 2020 to 2022), experimental as well as non-experimental exposure to screens, a broader array of screen types and platforms, more specific examination of vocabulary outcomes, and methodological moderators. Our goal is to provide additional insights into the conditions under which screen use may enrich or hamper vocabulary development. We proceed by laying out the arguments for potential moderators of the impact of screen use.

Potential moderators

Child age and gender

A couple of factors suggest that the opportunity costs of screen use may be more severe before age 3 years. First, evidence suggests that highly scaffolded interactions, which can be displaced by screens, are most critically important for infants and toddlers (Hassinger-Das et al., 2020). Second, a robust literature suggests the presence of a video deficit for early learning, such that young children are less likely to learn words from screen media than from in-person demonstrations. Such a video deficit is due to a range of factors, including younger children's limited cognitive maturation and experience with screen media (Barr, 2013; Mares & Sivakumar, 2014; Troseth et al., 2019). All of these barriers reduce the likelihood that young children generalize new vocabulary from screen media to real life.

Research suggests that children younger than 36 months old are especially unlikely to learn new information, including new words, from video compared to a real-life demonstration (Anderson & Pempek, 2005; Strouse & Samson, 2021). While few empirical studies have examined learning from screen media across a continuous age range, meta-analytic methods can maximize the value of these limited data and shed light on whether there is an abrupt, qualitative increase in learning from screens around the third birthday or a more gradual, quantitative increase across infancy and early childhood. Thus, the current study tested both accounts, examining a threshold effect of age by splitting our sample at 36 months and a monotonic effect of age across the full range of the sample by treating age as a continuous variable.

Additionally, given that Madigan et al. (2020) found that boys (vs. girls) showed greater gains in vocabulary associated with co-viewing with a caregiver, we examined child gender as a moderator. We could not examine family-level characteristics (e.g., socioeconomic status [SES]) as potential moderators, given limitations/inconsistencies in descriptions of study samples.

Media characteristics

The opportunity costs versus benefits of screen media use for vocabulary development may vary by media platform, in part because of the types of interactivity afforded by those platforms. Video chat can provide users with many of the same real-time, contingent turn-taking and feedback as in-person interactions (Myers et al., 2017; Roseberry et al., 2014). Apps and other interactive devices, including some e-books, can elicit turn-taking and provide feedback, but the feedback is likely to be more constrained than in a live interaction, and interactive media features may vary in the degree to which they support versus distract from word learning (Furenes et al., 2021). At the further end of the spectrum, some audiovisual narratives (e.g., TV shows) may mimic turn-taking by posing questions, taking pauses, or offering directives—so-called "pseudo-contingency"—but whatever feedback is given is not contingent on the child's response.

At the same time, platform tends to coincide with particular types of content structures. For example, television/video and e-books are more likely than games/apps or video chat to expose children to narratives, which may be particularly important for word learning (Linebarger & Vaala, 2010). E-books tend to be text-based even when they offer pictures and animation, whereas other platforms lend themselves primarily to audiovisual content. While these boundaries are not impermeable (e.g., a grandparent uses video chat to read narrative stories to their grandchild), it suggests that there may be platform differences even beyond issues of particular media features.

Regardless of platform, it seems crucial to consider the source of media: Content created by researchers for the purpose of studying word-learning could have different effects than professionally-created media. Additionally, within professionally-created media, child-directed content specifically designed to support early informal learning may provide the rich linguistic input and varied verbal environments (Schmidt et al., 2009) and types of visual aids considered critical to vocabulary learning (Paivio, 1986, 2007). Other content, designed more for entertainment, may not yield benefits that outweigh the opportunity costs of use.

Given these complexities, we examined whether associations between media use and vocabulary varied by platform (grouped broadly as TV/video, e-books, games/apps, and video chat), by interactivity, by source of media, and by educational intent.

Vocabulary measurement

Screen-based learning has been found to vary across different tasks, depending on task complexity, measurement sensitivity, and knowledge domain (for a review, see Kirkorian, 2018). For instance, there may be different media effects for measures of program-specific vocabulary versus overall vocabulary size. The former captures direct learning of words presented in media content, while the latter captures more general vocabulary development, perhaps resulting from cumulative exposure to educational media over time or increased ability to learn words in the environment (i.e., "learning to learn").

Similarly, children's performance varies for expressive vocabulary (i.e., word production) versus receptive vocabulary (i.e., word comprehension). Expressive vocabulary measurement is believed to represent a more precise association in memory between the word label and its corresponding semantics (Sénéchal & Cornell, 1993; Stahl, 1999), whereas receptive vocabulary measurement may rely on the visual context in which the target word is heard (Sénéchal, 1997). Given the evidence that different media features foster expressive and receptive vocabulary differently, and the mixed findings on the differential effects of screen media on these two types of vocabulary (e.g., Vatalaro et al., 2018; Verhallen & Bus, 2010), we examine vocabulary type as a potential moderator of media effects.

Other methodological differences

Finally, mixed findings about media and vocabulary development may be due to other methodological differences across studies (Krcmar, 2011). The relation between word learning and screen media exposure has been explored with multiple research paradigms that typically vary in method, media stimuli, measurement, environment, and study quality. For instance, experimental studies tend to use researcher-created stimuli, often in a lab setting, and the researchers assess acquisition of the specific, novel words taught in the session (e.g., Myers et al., 2017; Strouse et al., 2018). In contrast, correlational studies often focus on naturalistic exposure to commercial media, and the researchers tend to assess general vocabulary that may not have been featured in the media content (e.g., Hudon et al., 2013; Schmidt et al., 2009). Accordingly, we examined associations between media exposure and vocabulary separately for experimental and correlational research. Further, we examined potential methodological moderators, such as dosage for the treatment group (experimental studies), the context in which data were collected, and indicators of study quality.

The current study

As reviewed here, screen media have been found to hinder, benefit, or not influence children's vocabulary. Given the large body of relevant studies from a wide range of contexts, these inconsistencies may reflect the lack of a unified operational definition of screen media exposure, widely varying vocabulary measures, and/or disparate research methods and manipulations across studies. Examining these factors may help explain the mixed findings and shed some light on potential mechanisms underlying the influence of these factors on vocabulary learning in the context of screen media. Therefore, the present study aims at acknowledging the disparities and clarifying the ambiguities by clearly defining the concepts and variables studied, setting strict inclusion criteria, and considering test-related factors as confounding moderators. These analyses were primarily exploratory insofar as: (1) we did not know which of the competing hypotheses would be supported in mixed findings in past research and (2) it was not clear until coding was completed which moderators could be validly included.

The current meta-analytic study investigated associations between early screen media use (birth to 6 years) and children's vocabulary. We examined early media use given the critical importance of identifying early experiences that may lead to differences in vocabulary size, which in turn predicts a wide range of later outcomes (Dickinson & Porche, 2011; Pace et al., 2019; Snow et al., 1998). We focused on unaided foreground screen media use and first-language vocabulary acquisition in typically developing children.

METHOD

Search strategies

Relevant studies were collected in four steps. First, we searched five major databases for journal articles, technical reports, book chapters, and theses/dissertations from the year each database started until Sept 1st, 2022. The databases included Web of Science, Ebsco, PsycINFO, Communication and Mass Media Complete, and WorldCat Theses/Dissertations, using one search term from each of two categories (247 pairs in all):

- Media: media*, screen*, touchscreen*, touch-screen*, touch screen*, video*, tv, telev*, dvd, program*, sesame street (given the relatively large number of studies on this program), game*, comput*, app*, video-chat, video chat, e-book, ebook, electronic book
- Outcome: vocab*, word*, language*, lingu*, literacy, CDI, PPVT, EVT, ROWPVT, EOWPVT, cognit*, learn*, achieve*

Second, we inspected the reference lists of all studies that met our inclusion criteria and the reference lists for relevant review articles and book chapters. Third, we conducted a citation search in Google Scholar, capturing publications that cited each included study. Last, we collected unpublished work by searching relevant websites (e.g., PBSKids, Corporation for Public Broadcasting) for technical reports of program evaluations, soliciting unpublished datasets through professional listservs (Cognitive Development Society, American Psychological Association, International Congress on Infant Studies), and contacting authors whose work was included in our dataset or cited in our literature review multiple times. See Figure 1 for details of the search process.

Inclusion and exclusion criteria

Studies had to meet four criteria to be included. First, the focus of this meta-analysis was on children's foreground screen media use rather than incidental or background exposure. Second, experimental research had to include a manipulation of screen media exposure; correlational research had to include a measure of screen media exposure. Third, the studies had to assess children's receptive or expressive vocabulary. The vocabulary measure could capture study-specific word-learning (e.g., object selection or looking time in an object-labeling task) or global vocabulary using a standardized test (e.g., Peabody Picture Vocabulary Test) or parent-report questionnaire (e.g., MacArthur Communicative Development Inventory). Fourth, we only included studies/conditions with child participants aged 0 and 6 years old at the time of media exposure. Longitudinal studies in which media exposure was measured at or before 6 years of age were included, even if the vocabulary assessment was conducted after 6 years of age, consistent with other recent meta-analyses of media effects (e.g., Madigan et al., 2020; Strouse & Samson, 2021).

In addition, three exclusion criteria were applied. First, given our focus on unaided media use, we excluded studies or conditions in which a person was present in the room with the child to assist or guide the child's use of the screen media. Specifically, we excluded studies that involved adult explanations and discussion (e.g., parent or teacher active mediation) or peer scaffolding (e.g., Neuman & Kaefer, 2013) that were intended to support children's learning from the screen. For example, we excluded conditions that involved reading buddy programs (e.g., Martha Speaks Reading Buddies), which pair younger and older children to read together (e.g., Silverman, Kim, et al., 2017; Silverman, Martin-Beltran, et al., 2017), and adult-involved multimedia programs (e.g., World of Words), designed for adults to engage children with interactive techniques (e.g., Neuman & Kaefer, 2013). In other studies, we excluded comparison

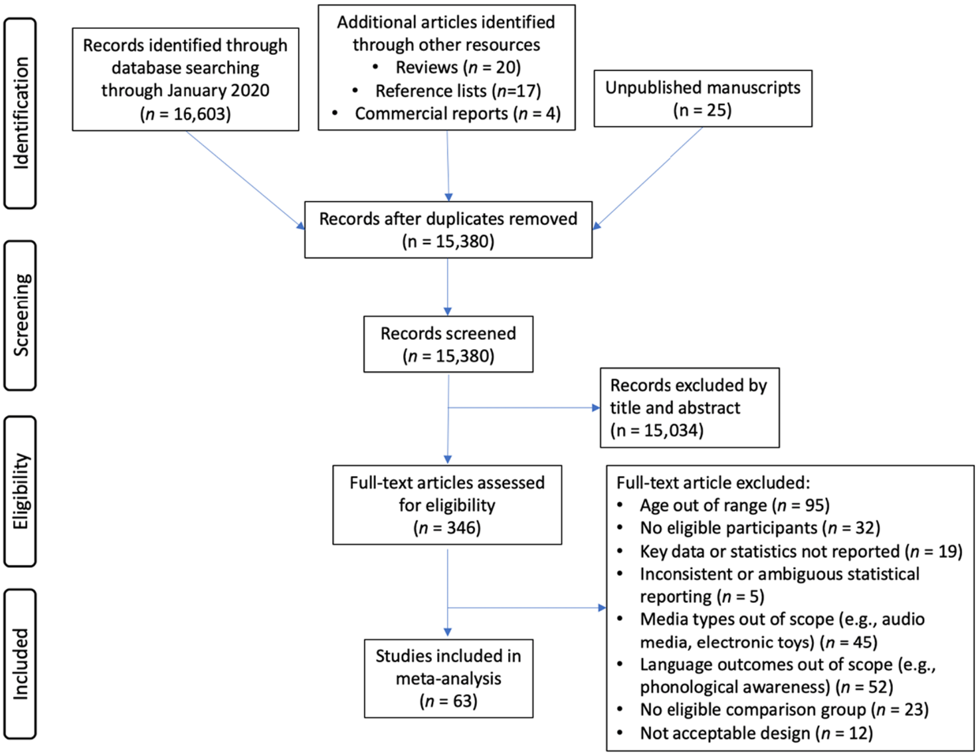

FIGURE 1 PRISMA flow diagram depicting literature search and screening.

conditions that involved adults reading and explaining printed books (e.g., De Jong & Bus, 2002; Segal-Drori et al., 2010). Second, given different effects of screen media on the development of first language and foreign language found in preschool age children (e.g., Aslanabadi & Rasouli, 2013; Silverman, Martin-Beltran, et al., 2017), research that examines second language learning was not considered. For studies involving first-language and second-language learning, only the data from first-language learner samples were included (e.g., Hudon et al., 2013). Lastly, effects were excluded if they focused on atypical development, such as interventions designed to ameliorate delays in language and literacy skills (e.g., Segers et al., 2006; Verhallen et al., 2006).

General coding procedure

We coded each study for four types of information: (1) bibliographic information (i.e., authors, year of publication, and title of study); (2) basic study information (i.e., study scale, research method, independence relation among subgroups within studies); (3) characteristics related to the potential moderators (e.g., child age and gender, media platform, vocabulary type); and (4) quantitative information for the calculation of effect sizes (e.g., group sample sizes, means, and standard deviations for experimental research; correlation or regression coefficients for correlational research). To compute inter-rater agreement, 20 articles (i.e., 32%) were selected at random and coded by two independent coders. Satisfactory inter-rater agreement was achieved with an average intraclass correlation of .94 for continuous variables and an average Kappa of .93 for categorical variables. All coding discrepancies were discussed and resolved by the first and second authors.

Moderators

We considered four types of moderators: (1) study quality and design; (2) child characteristics; (3) media characteristics; and (4) vocabulary measurement.

Study quality and design

Research method captured whether the study was experimental or correlational (κ=1.00). Experimental designs include between-subject design, within-subject design, or mixed design. Correlational studies included cross-sectional and longitudinal research. *Publication status* captured whether the study was published in a peer-reviewed journal (κ=1.00). *Publication year* was the year of publication (ICC r=1.00). Measurement design captured whether the study measures were collected at a single time point or across multiple points (κ=1.00). Following Ulferts et al. (2019), *child covariates* captured whether the analysis controlled for any child characteristics (κ=.91), while *context covariates* captured whether the analysis controlled for any family, household, or school characteristics (κ=.98). *Test environment* captured whether vocabulary assessment occurred at home, in school, in a research lab, or at another location (κ=.94). Finally, *treatment dosage* captured whether children had a single exposure or multiple exposures to experimental stimuli (e.g., a video designed to teach specific words; κ=.84). Note that *treatment dosage* was not included as a moderator for correlational studies, because (1) there was no treatment involved in correlational studies and (2) the dosage of screen media exposure is analyzed as the main predictor (not a moderator). Put differently, dosage was captured as the main effect for correlational studies and a moderator in experimental studies.

Child characteristics

Child characteristics included age and gender. *Child age* at the time of media exposure was coded in two ways: as a continuous variable (in months; ICC r=.90) and as a dichotomous variable (less than 36 months vs. more than or equal to 36 months). For studies that reported age with a range or did not specify the month, we used the midpoint of the range. For example, "24–30 months old" and "2 years old" were coded as 27 and 30 months, respectively. *Child gender* was coded as the proportion of the sample that was female (ICC r=1.00).

We would have liked to consider a wider range of child and family characteristics. Indeed, we coded our data set for child race/ethnicity and four indicators of SES, including household income, parent education, parent marital status, and parent occupational status. However, this information was reported infrequently and inconsistently across studies. Thus, there was not sufficient information coded in the current dataset to conduct robust moderator analyses. A qualitative overview of reported participant socio-demongraphic information for each study can be found in Supporting Information S1 and S2.

Media characteristics

Media platform was coded in four categories (κ=.81): TV/video referred to narrative-based, observational (i.e., noninteractive) screen media showing moving images (e.g., TV programs, DVDs, other videos). E-books referred to text/audio with static images, often designed for literacy engagement. Games/apps included digital games or apps that incorporated interactivity and featured entertainment. Video chat referred to interacting in real-time with a live on-screen social partner through video chat technology (e.g., FaceTime, live closed-circuit video).

Media source captured whether the media exposure involved professionally produced products available on the market, researcher-created materials that were designed for the study, or researcher-edited materials in which the researcher edited professionally produced material for the study (κ=.95).

Media interactivity referred to media that intentionally seek children's input or in which the input determines how the content is displayed on the screen (κ=.83). We defined and coded interactivity in a stringent way that entails the contingent and reciprocal response between the child and screen media. In this sense, pseudo-contingency typically featured in TV shows (e.g., *Blue's Clues*) that cannot adapt to input from the child was not coded as interactive. In our sample, interactive media included media games or touchscreen apps (e.g., Russo-Johnson et al., 2017) and e-books with interactive hotspots, icons, or mini-games (e.g., Smeets & Bus, 2015) that allowed interactivity by prompting children's inputs or giving feedback as well as video chatting with another person in real time (e.g., Myers et al., 2017). For example, e-books with forward/backward buttons were coded as interactive, since they are designed for children to control the content display and reading flow (e.g., Korat & Shamir, 2012). E-books that were simply on-screen pictures of printed storybooks, accompanied by a voiceover of the text (e.g., the static e-book condition in Smeets & Bus, 2015) were not coded as interactive, nor were storybooks on video without built-in functions to prompt reciprocal contingent interactions with children (e.g., regular video conditions in Strouse et al., 2013).

Media content captured whether or not the content was identified by the researcher as educational, in that it was specifically designed to foster children's learning or development (e.g., *Sesame Street, Baby Einstein, Your Baby Can Read*; κ=.94). Studies where the researchers did not identify the content as educational may or may not have contained educational content and was coded as "unspecified". Note that this moderator did not apply to experimental studies, because all experimental studies, by definition, tested the effect of media designed to teach one or more words. That is, the main effect of media exposure in experiments is itself a test of content with educational intent (i.e., treatment vs. control). Therefore, this code was used as a moderator for correlational studies only.

Vocabulary measurement

Vocabulary type distinguished between tests of receptive versus expressive vocabulary (κ=.98). *Vocabulary specificity* captured whether the dependent variable was direct learning of media-specific vocabulary (e.g., learning a word presented in experimental stimuli) or a global estimate of general vocabulary size, regardless of whether words on the vocabulary test appeared in the media (e.g., using the Peabody Picture Vocabulary Test or the MacArthur-Bates Communicative Development Inventories; κ=.89). *Vocabulary source* captured whether the dependent variable was measured through direct assessment of children or through parent report ($\kappa = .90$).

Coding and calculation of effect sizes

Experimental effect sizes

In experimental studies, the effect size of screen media exposure on vocabulary outcomes denotes the degree to which vocabulary outcomes differed between children who were exposed to the program-specific vocabulary via screen media and those who were not. Since children's vocabulary performance is measured as a numerical variable in most experiments included in this analysis, Cohen's *d* was computed using mean and standard deviation of vocabulary performance to quantify the standardized difference in vocabulary outcomes between children exposed to and not exposed to certain screen media.

Three types of comparison were coded for the calculation of experimental effect sizes, depending on different experimental designs. For *between-subject* studies, we coded the comparison between a control condition and a treatment condition in which the participants were exposed to screen media. In treatment conditions, children viewed screen media stimuli that presented the target word(s). Control conditions could include participants who were not exposed to any media (e.g., De Jong & Bus, 2002) or those who were exposed to equivalent media that did not contain the target words (e.g., Rice & Woodsmall, 1988; Segers & Verhoeven, 2003). The effect size was computed based on the mean and standard deviation of numerical vocabulary performance in the treatment and control condition, following the formula proposed by Borenstein et al. (2009, pp. 26–27).

For *within-subject* designs, two types of comparison were coded. Most studies involved a pre-post design, comparing each child's vocabulary measured before versus after exposure to a treatment stimulus. Effect size was calculated based on the mean and standard deviation of vocabulary performance at pre-test and post-test, following the procedure outlined by Borenstein et al. (2009, p. 29). In rare cases (e.g., Roseberry et al., 2014), the comparison was between children's learning of target words (i.e., program-specific vocabulary) and that of control words (i.e., words not presented in the screen media stimuli), and the effect size was computed based on the mean and standard deviation of performance on target and control words, using the same procedure. We used $r = .5$ as the repeated-measure correlation based on the adjustment correction of Strouse and Samson (2021).

For *mixed* design studies, both the pre- and post-test performance in both the treatment and control condition were coded. The effect size was calculated based on the comparison between pre-post gains for the treatment group versus control group, following a *difference-in-difference* approach (Morris, 2008, equations 8, 9, and 24) and consistently using $r = .5$ as the repeated-measure correlation.

For studies that did not have a valid baseline control group, we assumed a chance-level performance for the control group, as long as children were tested on novel words (whether unfamiliar or nonwords, e.g., Kirkorian et al., 2016). Specifically, we modeled the control performance using a random binomial distribution, because the experiments included in this analysis measured vocabulary outcome by recording the number of successful trials, each with a binary outcome (i.e., successful vs. unsuccessful). The mean and standard deviation of the control performance were thereby computed based on the calculation of expected value and variance of binomial distribution.

Furthermore, while most experimental studies measured vocabulary as an ordinal variable for effect size calculation using mean and standard deviation as described above, a few had binary vocabulary outcomes, such as the proportion of participants who passed or failed the test (e.g., Kirkorian et al., 2016). In this circumstance, we converted the proportions to d_{Cox}, a logistic transformation of the odds-ratio, as the effect-size index for Cohen's *d* (Chinn, 2000; Sánchez-Meca et al., 2003).

Correlational effect sizes

In correlational studies, effects represent the strength of the correlation between the amount of screen media exposure (i.e., predictor variable) and children's vocabulary (i.e., outcome variable). We coded two commonly used metrics of correlational effect sizes: correlation coefficient (r) and standardized regression coefficient (β). The use of these two metrics remains controversial. The equivalents of bivariate r (e.g., correlation coefficient, uncontrolled regression coefficients) have been widely used in meta-analyses for correlational studies. However, given concerns (e.g., Ferguson, 2015; Furuya-Kanamori & Doi, 2016) that this approach yields spurious estimates of true effects by not adjusting for confounding variables, recent meta-analyses, including those in behavioral sciences, have used standardized regression coefficients (i.e., beta coefficient β) from multivariate analyses with covariates included (e.g., Becker & Wu, 2007; Peterson & Brown, 2005). Meanwhile, some researchers (e.g., Rothstein & Bushman, 2015; Valkenburg, 2015) question whether standardized regression coefficients from different studies can be meaningfully aggregated and compared, given that they tend to be based on models with different covariates (Lipsey & Wilson, 2001). Given this ongoing debate, we considered both metrics while including a sensitivity analysis to check the possible influence of using different metrics on the overall estimate. We converted β to r, following Peterson and

Brown's (2005, p. 171) formula, which has been shown to be reliable if r ranges from −.50 to .50 (Bowman, 2012; Hunter & Schmidt, 2004; Peterson & Brown, 2005).

To combine experimental and correlational effect sizes, we first converted Cohen's d from experimental studies to r using Rosenthal's (1994, p. 48) formula. Next, we used Fisher's transformation to convert r to z-statistic prior to the aggregation to overall effect size and transformed back for interpretative purposes (Bowman, 2012, equations 1 and 2).

Missing data strategies

Twenty-five studies were missing statistics needed for effect size calculation. We scanned these papers for other statistics (e.g., t-tests and p-values) that could be converted to standardized mean differences following procedures in Harrer et al. (2021); however, none of the studies contained sufficient information to perform the conversions. As a final attempt to retain these studies, the authors were contacted via email. Six articles were retained for a 24% response rate.

Some articles lacked the information needed to code moderators. Where possible, estimated values were coded and the estimating criteria were consistently applied to similar missing cases. For example, Rice et al. (1990) described participants as within 3 months of their third birthdays, so we used the midpoint of the range to estimate the average age (i.e., 36 months). However, most moderators could not be estimated (e.g., environment in which children were exposed to the screen media) and were coded as missing if the relevant information was not provided.

Analytic approach

Examination of outliers

Since the presence of outliers and influential data points may reduce the validity and robustness of the conclusion drawn from a meta-analysis (Viechtbauer & Cheung, 2010), all effect sizes were inspected for influential outliers. Effects sizes with an absolute standardized score (i.e., z value) higher than 3.29 (Tabachnick et al., 2007) or an absolute studentized deleted residual higher than 2 (Nikkelen et al., 2014) were considered as influential. Cook's distances and DFBETAS values were also calculated as a cross-check following Viechtbauer and Cheung's procedure (2010) developed for random effects and mixed-effects meta-analytic models. Identified outliers were windsorized into values of .01 beyond the largest non-outlier effect size (see Results).

Multi-level modeling

Most studies yielded multiple effect sizes. To maximize the inclusion of individual differences as moderators, we coded separate effect sizes for subgroups (e.g., different age groups) whenever possible, instead of coding the overall effect size averaged across all subgroups. Thus, we used a three-level random-effects model (Cheung, 2014; Van den Noortgate et al., 2013) to account for two sources of dependence: dependence among subgroups within studies (e.g., multiple experiments reported in the same paper) and dependence among effect sizes within subgroups (e.g., multiple tests using the same sample). A series of multilevel random-effect models were fitted using the rma.mv() function in the metafor R package (Viechtbauer, 2010). Each effect size was weighted by the inverse of its variance (Card, 2012; Cooper, 2009). We further used the robust() function to obtain cluster-robust standard error estimates based on an estimate of the variance–covariance matrix. A compound symmetric variance–covariance structure was chosen to avoid making any assumptions about the relations among the random effects.

Heterogeneity and moderator analyses

Given the significant degree of heterogeneity suggested by the homogeneity statistic and between-group variance (see Results), moderator analyses were conducted to test whether the variance among individual effect sizes can be explained by the group-level moderators. Meta-regression models were run for continuous moderators (Viechtbauer, 2007), while the procedures outlined by Viechtbauer (2015) were used for categorical moderators. Separate models were run with each moderator variable. We also conducted sensitivity analyses to test moderators within the two research methods (experimental, correlational) and to compare experimental and correlational effect sizes within each moderator category.

Examination of potential quality and publication bias

Moderator analyses of study quality characteristics were conducted to assess risk of bias. To reduce bias of overestimating the true effect sizes due to missing reports (Borenstein et al., 2009), publication bias was evaluated by inspecting funnel plots and Egger's regression test. This is particularly relevant in the current project given the number of unpublished technical reports for commercial screen media content.

RESULTS

Our sample included 44 experimental studies (including one quasi-experimental study) and 19 correlational studies published between January 1988 and September 2022. Of these 63 studies, the majority (n=48) assessed English vocabulary, with fewer studies assessing Dutch (n=4), Hebrew (n=4), French (n=2), Chinese (n=2), Turkish (n=1) and Arabic (n=2). Among samples with information available about race/ethnicity (51%) and sex/gender (73%), most were majority White/Non-Hispanic and bewteen 40%–60% female. There was a total of 266 valid effect sizes and 11,413 participants involved, ranging in age from 0.90 to 6.58 years. See Supporting Information for coded moderators (S3 and S4) and effect sizes (S5 and S6) for each subsample. Across all samples, the average age at the time of screen media exposure was 40.44 months, and gender distribution was on average 48% female and 52% male.

Overall weighted average effect size

The overall effect size was small but positive and significant (r=.23, 95% CI [.17, .30], p<.001). Among the 266 effect sizes that emerged from the dataset, 40 were large (>0.5), 61 were medium (>0.3), 77 were small (>0.1), and 78 were very small (<0.1), according to Cohen's (1988) standard. Four effect sizes from two studies were identified as influential outliers and were windsorized. However, the overall effect size was quantitatively similar and qualitatively the same before windsorizing outliers (r=.23, 95% CI [.17, .30], p<.001). The effect sizes displayed a largely symmetric distribution which centered to the right of zero on the x axis (see Figure 2). The forest plot with 95% confidence interval for each effect size is shown in Supporting Information S5 and S6.

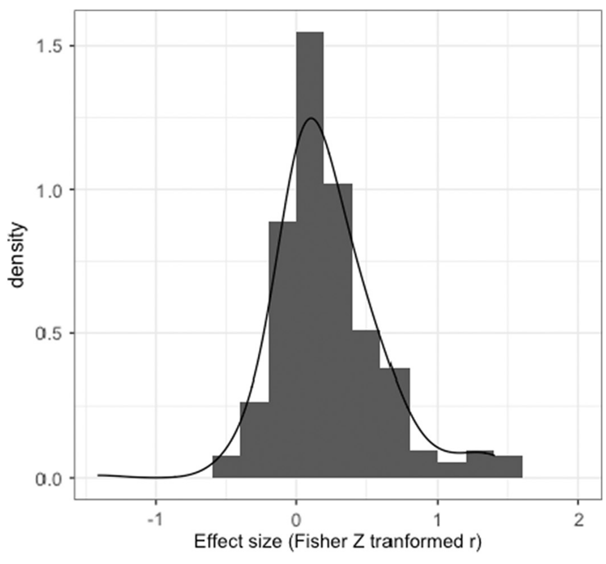

FIGURE 2 Distribution of effect sizes included in the meta-analysis.

The homogeneity statistic (Q_B=2121.93) indicated a significant degree of heterogeneity in the effect sizes (p<.001), with high variance at the subgroup level in addition to random error (I^2=91%). Given this, moderator analyses were conducted to examine factors that might explain this variance.

Moderator analyses

Table 1 shows the results of all moderator analyses and sensitivity checks. Sensitivity checks include subsample moderator analyses for each of the two research methods (experimental, correlational) and individual effect size estimates for each level of categorical moderators overall and within the two research methods.

Study quality and design

Research method was a significant moderator (k = 266, Q_B = 13.19, p < .001). The average effect size was larger for experimental studies than correlational studies (b=.23, 95% CI [.11, .35], p < .001). There was a small-to-medium effect for experimental studies (r=.30, 95% CI [.22, .38], p<.001), while the effect for correlational studies was close to zero (r=.07, 95% CI [−.01, .16], p=.066). There were no other significant moderators based on study quality or design. The size of media effects did not differ by publication status (peer-reviewed vs. nonpeer-reviewed study; k=266, Q_B=1.09, p=.164), publication year (k=266, Q_B=0.04, p=.836), measurement design (single vs. multiple time points; k=263, Q_B=1.27, p=.268), inclusion of child-level (k=83, Q_B=0, p=.937) or context-level (k=83, Q_B=0.03, p=.864) covariates, test environment (k=233, Q_B=8.86, p=.115), or treatment dosage (k=177, Q_B=0.15, p=.696). See Table 1 for overall moderator analyses and sensitivity checks within subsamples.

Child characteristics

Continuous age

We first tested average sample age as a continuous moderator. A meta-regression analysis revealed that age significantly moderated the effect of screen media exposure on vocabulary, though it only accounted for a small portion of the heterogeneity (k=266, Q_B=5.00, p=.025). Effect sizes increased with age (b=.004, 95% CI [.001, .007], p=.025), indicating that the positive relation between screen media and vocabulary grew stronger with age (see Figure 3). Sensitivity analyses showed that the effect of average sample age was significant in experimental studies (b=.005, 95% CI [0, .009], p=.039) but not

TABLE 1 Moderator and sensitivity analyses across and within research method for categorical moderators.

	All studies				Experimental subsample				Correlational subsample			
	Moderator analysis		Effects by category		Moderator analysis		Effects by category		Moderator analysis		Effects by category	
	k	Q_B	*k*	*r*	*k*	Q_B	*k*	*r*	*k*	Q_B	*k*	*r*
Study quality and design												
Research method	266	13.19***			—	—			—	—		
Experimental			183	.30***,a			—	—			—	—
Correlational			83	.07b			—	—			—	—
Publication status	266	1.09			183	1.20			83	0.69		
Peer-reviewed			216	.25***			156	.32***			60	.07
Other			50	.17**			27	.19**			23	.16
Publication year	266	0.04			183	0.01			83	−0.01		
Measurement design	263	1.23			183	0.10			80	0.01		
One single time point			181	.21***			113	.30***			68	.07
Multiple time points			82	.25**			70	.30***			12	−.03
Child covariates	83	0.00			—	—			83	0.00		
Included			31	.08			—	—			31	.08
Not included			52	.07			—	—			52	.07
Context covariates	83	0.03			—	—			83	0.03		
Included			28	.06			—	—			28	.06
Not included			55	.08			—	—			55	.08
Test environment	233	8.86			168	3.01			75	1.06		
Lab			87	.28***			83	.31***			4	.02
Home			71	.12†			22	.16			49	.11
School			75	.27***			63	.32***			12	.03
Treatment dosage	177	0.15			177	0.15			—	—		
One single exposure			86	.31***			86	.31***			—	—
Multiple exposures			91	.29***			91	.29***			—	—
Child characteristics												
Continuous age (months)	266	5.00*			183	5.08*			83	0.72		
Age group	266	6.70**			183	6.46*			83	0.21		
Below 36 months			112	.15**,a			69	.20**,a			43	.07
At or above 36 months			154	.28***,b			114	.37***,b			40	.06

TABLE 1 (Continued)

	All studies				Experimental subsample				Correlational subsample			
	Moderator analysis		Effects by category		Moderator analysis		Effects by category		Moderator analysis		Effects by category	
	k	Q_B	*k*	*r*	*k*	Q_B	*k*	*r*	*k*	Q_B	*k*	*r*
Female (%)	157	1.15			130	1.85			27	0.11		
Media characteristics												
Media platform	264	17.46***			183	13.13**			—	—		
E-books			75	.38***[,a]			73	.40***[,a]			2	−.02
TV/video			139	.15***[,b]			73	.20**[,b]			66	.08
Games/apps			38	.19**[,b]			27	.25**[,b]			11	.09
Video chat			12	.46[†]			10	.66*			2	−.02
Media source	258	9.69*			175	0.78			—	—		
Professionally-produced			162	.16***[,a]			79	.26***			83	.07
Researcher-created			83	.37***[,b]			83	.36***			0	—
Researcher-edited			13	.35**[,b]			13	.35**			0	—
Media interactivity	266	5.00*			183	3.02*			83	0.09		
Interactive			112	.35***[,a]			99	.39***[,a]			13	.07
Non-interactive			154	.17**[,b]			84	.25***[,b]			70	.08
Media content	55	8.18**			—	—			55	8.18**		
Educational			21	.17*[,a]			—	—			21	.17*[,a]
Unspecified			34	.01[b]			—	—			34	.01[b]
Vocabulary measurement												
Vocabulary specificity	265	29.55***			182	11.14***			—	—		
Program-specific			162	.35***[,a]			162	.35***[,a]			0	—
General			103	.10*[,b]			20	.18[b]			83	.07
Vocabulary type	255	6.99**			182	7.22**			73	1.91		
Receptive			191	.24***[,a]			143	.26***[,a]			48	.02
Expressive			64	.36**[,b]			39	.44***[,b]			25	.10
Vocabulary source	80	2.27			—	—			80	2.27		
Direct assessment			49	.03			—	—			49	.03
Parent-report			31	.14*			—	—			31	.14*

Note: Models were fit for the entire sample ("All Studies") and within each subsample based on research method ("Experimental", "Correlational"). Significant Q_B values indicate significant degree of variance accounted by the moderator. Significant *r* values indicate significant difference from zero (i.e., significant media effect) for a given category. Different superscripts (i.e., a vs. b) within a categorical moderator indicate significant differences between effect sizes for the corresponding categories.

[†]$p < .08$.

*$p < .05$; **$p < .01$; ***$p < .001$.

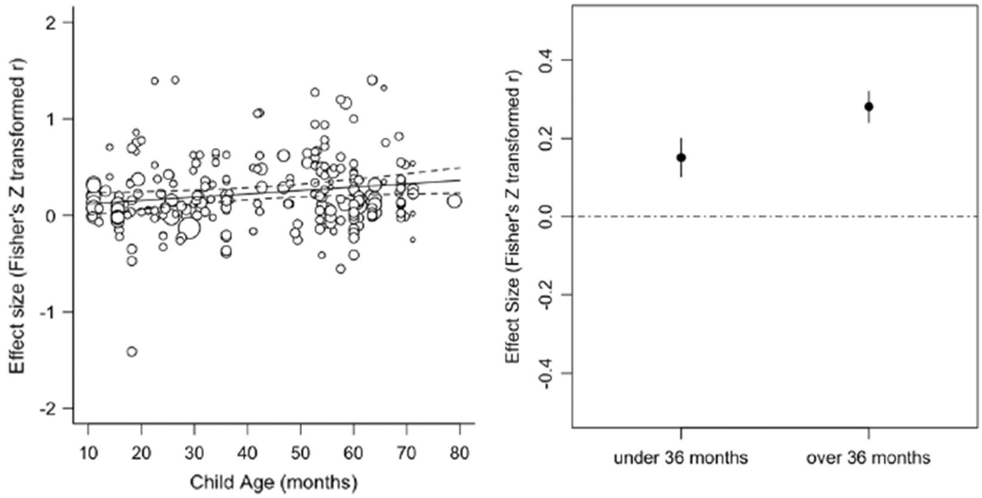

FIGURE 3 Age moderation: meta-regression (left) and age group comparison to chance (right). The size of the circle in the left plot is proportional to study weight (i.e., inverse of variance) in the aggregated effect size. The right plot contains the weighted average effect sizes by age group and the 95% confidence intervals.

in correlational studies ($b=-.001$, 95% CI [−.003, .001], $p=.154$). Note that the average age was higher in experimental studies ($M_{age}=49$ months) than in correlational studies ($M_{age}=33$ months; $t(158)=3.35$, $p<.001$), which might partly account for the overall effect of continuous age across all studies.

Age group

We also tested average sample age as a categorical moderator, comparing effect sizes for studies with an average age above versus below the threshold of 36 months. Age group was a significant moderator across all studies ($k=270$, $Q_B=6.70$, $p=.010$). The average effect size for studies with an average participant age of 36 months or older was significantly larger than for studies with an average participant age below 36 months ($b=.12$, 95% CI [.03, .22], $p=.013$). This effect held within experimental studies ($b=.19$, 95% CI [.04, .36], $p=.015$), but not within correlational studies ($b=-.02$, 95% CI [−.09, .05], $p=.510$). As shown in Table 1; Figure 3, even though the average effect size was smaller for younger children than older children, the weighted average effect size was significantly larger than zero for both age groups (younger: $r=.15$, 95% CI [.06, .23], $p=.002$; older: $r=.28$, 95% CI [.20, .37], $p<.001$). Sensitivity checks indicated that this was also true within experimental studies but not within correlational studies (see Table 1).

Gender

The percentage of female participants included in each study was not a significant moderator ($k=157$, $Q_B=1.15$, $p=.283$). This result was consistent across the research methods (experimental: $k=130$, $Q_B=1.85$, $p=.173$; correlational: $k=27$, $Q_B=0.11$, $p=.743$).

Media characteristics

Media platform

The effect sizes of screen media exposure on vocabulary learning were significantly moderated by media platform ($k=264$, $Q_B=17.46$, $p<.001$). Specifically, the average effect size was larger for e-books than TV/video ($b=.25$, 95% CI [.12, .38], $p<.001$) and games/apps video ($b=.21$, 95% CI [.08, .34], $p=.002$). Sensitivity analyses indicated that the average media effect was positive and significant for three of the four platforms (e-books: $b=.38$, 95% CI [.27, .50], $p<.001$; TV/video: $b=.15$, 95% CI [.07, .22], $p<.001$; games/apps: $b=.19$, 95% CI [.10, .29], $p=.001$). Video chat was the only platform with a non-significant media effect across all studies ($b=.46$, 95% CI [−.04, .96], $p=.063$), despite having the largest average effect size (see Table 1).

The media platform moderator effect was robust in experimental studies ($k=183$, $Q_B=13.13$, $p=.001$), in which the average effect size was larger for e-books than for TV/video ($b=.22$, 95% CI [.06, .38], $p=.008$) and games/apps video ($b=.18$, 95% CI [.03, .34], $p=.019$). In the experimental subsample, the average effect of media was positive and significant for all platforms (e-books: $b=.40$, 95% CI [.28, .52], $p<.001$; TV/video: $b=.20$, 95% CI [.09, .31], $p=.002$; games/apps: $b=.25$, 95% CI [.11, .38], $p=.004$; video chat: $b=.66$, 95% CI [.03, 1.29], $p=.044$). There were too few correlational studies for most platform categories to examine media platform as a moderator in this subsample.

Media source

The source of media stimuli was a significant moderator ($k=258$, $Q_B=9.69$, $p=.008$). The average effect size was smaller for professionally-produced media compared to researcher-created (b = .20, 95% CI [.06, .34], p = .005) or researcher-edited (b = .14, 95% CI [0, .28], p = .046) media. However, sensitivity checks indicated that this was not true within experimental studies (k = 175, Q_B = 1.52, p = .468). We could not examine media source as a moderator in the correlational subsample since none of the correlational studies involved lab-created or lab-edited media. Note that, as shown in Table 1, the average media effect was significantly different from zero within all media source categories in the full sample and in the experimental subsample. Additional analyses not shown in Table 1 revealed that, even among studies using professionally-produced media, experimental studies generated significantly larger effects than correlational studies (b = .18, CI [.03, .33], p = .019).

Media interactivity

Interactivity was a significant moderator (k = 270, Q_B = 5.00, p = .025). The effect size for exposure to interactive media was larger than the effect size for exposure to non-interactive media (b = .11, 95% CI [.01, .23], p = .025). This result was robust in experimental studies (b = .11, 95% CI [0, .24], p = .032) but not in correlational studies (b = −.02, 95% CI [−.04, .02], p = .231). Nonetheless, as shown in Table 1, both interactive and non-interactive media had significant effects across research method and within the experimental subsample.

Media content

As described earlier, we looked within the subsample of correlational studies to examine whether effects of exposure to TV/video varied by whether the content was explicitly designed to be educational (k = 55, Q_B = 8.18, p = .004). The average effect size for exposure to educational TV/video was larger than the average effect size for exposure to unspecified TV/video (b = .11, 95% CI [.03, .18], p = .004). In fact, as Table 1 shows, the average effect was significant only for educational TV/video (r = .17, 95% CI [.01, .34], p = .042); the average effect for exposure to unspecified TV/video was close to zero (r = .01, 95% CI [−.09, .12], p = .751).

Vocabulary measurement

Vocabulary specificity was a significant moderator (k = 265, Q_B = 29.55, p < .001). The effect size for program-specific words was larger than for children's general vocabulary (b = .30, 95% CI [.19, .42], p < .001). The moderator effect was robust in experimental studies (k = 182, Q_B = 11.14, p < .001), in which the average effect size was again larger for program-specific words than for general vocabulary (b = .32, 95% CI [.13, .51], p < .001). While the overall media effect was significantly larger than zero for both program-specific and general vocabulary, the average media effect was significant for program-specific vocabulary only in the experimental subsample (see Table 1). We could not examine vocabulary specificity as a moderator in the correlational subsample since none of the correlational studies measured program-specific vocabulary. Moreover, among studies measuring general vocabulary, effect sizes did not differ for experimental versus correlational studies (b = .07, CI [−.01, .16], p = .066).

Vocabulary type was also a significant moderator (k = 255, Q_B = 6.99, p = .038). The effect size for expressive vocabulary was larger as compared to receptive vocabulary (b = .14, 95% CI [.01, .27], p = .038). This result was robust in experimental studies (b = .20, 95% CI [.03, .37], p = .025), but not in correlational studies (b = .05, 95% CI [.03, .14], p = .214).

There was no moderating effect of vocabulary source (parent report vs. direct assessment; correlational: k = 80, Q_B = 2.27, p = .077). See Table 1 for moderator analyses and sensitivity checks overall and within subsamples.

Evaluation of publication bias

Diagnostic tests for evaluating publication bias in the context of multi-level modeling are still evolving. It remains debatable whether the trim and fill method is a reliable and informative indicator of publication bias in multi-level data sets (e.g., Coburn & Vevea, 2015; Van Assen et al., 2014). The fail-safe N method has been criticized for its statistical weakness and not providing a valid assessment of publication bias (e.g., Becker, 2005; McDaniel et al., 2006). Funnel plots, (together with the weighted Egger's test and the rank test) are frequently used to test the asymmetry of data distribution, but these are also not a trustworthy indicator of publication bias in multi-level datasets, because they assume independence among individual effect sizes, (e.g., Coburn & Vevea, 2015; Van Assen et al., 2014). However, funnel plots can be used to visually evaluate whether there is a pattern in the data (e.g., Verhoef et al., 2019). Figure 4 shows the funnel plot that depicts individual effect sizes (on the x axis) against their standard error (on the y axis). Visually, the plot appears largely symmetrical, with individual effect sizes distributed evenly on both sides of the vertical line (i.e., the estimated overall effect size). A few outliers at the bottom with relatively big variance and small sample size are from the same study. A sensitivity test was conducted by removing these outliers, and the results remained quantitatively similar and qualitatively the same (see Supporting Information S7).

To statistically test publication bias, we extended the Egger's test by modifying our multi-level model to include the standard error of individual effect sizes as a moderator, following Habeck and Schultz's (2015) procedure. The slope of the moderator did not significantly deviate from zero (b = .85, 95% CI [−.06, 1.76], p = .066), suggesting that the size of effects was not related with their

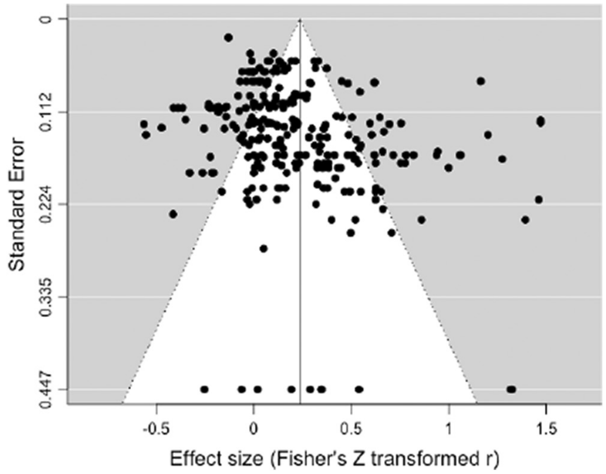

FIGURE 4 Funnel plot for individual effect sizes by their standard error.

standard errors, and therefore, there was no evidence of bias (Rothstein & Bushman, 2015). In addition, publication status (i.e., peer-reviewed vs. nonpeer-reviewed study) was not a significant moderator ($k=266$, $Q_B=1.09$, $p=.164$), indicating no evidence of publication bias. As a further sensitivity check, we compared fitted models with and without influential effect sizes as defined earlier in the method section; the results were robust to the inclusion of these influential outliers.

DISCUSSION

This meta-analysis provides a quantitative review of studies that have examined the influence of screen media on young children's vocabulary. The goal was to shed light on the mixed findings that have been reported in the empirical literature. Combining 63 studies including 11,413 children, this meta-analysis revealed an overall positive relation ($r=.23$) between screen media exposure at age 6 years or younger and concurrent or subsequent vocabulary. Notably, this association was observed even in the absence of adult involvement, despite the well-documented effect of adult mediation in helping young children interpret content and learn words from screens (Madigan et al., 2020; Strouse et al., 2018). However, the moderator analyses suggest this media effect varies by study design, child characteristics, media characteristics, and vocabulary measurements.

To summarize, the positive association between screen media exposure and vocabulary was largely limited to experimental contexts rather than "real world" settings, with stronger effects when researchers used content they had created, rather than extant content produced by professionals, and when they assessed learning of the specific words taught in the content. Non-experimental, "naturalistic" screen media use was not associated with vocabulary, with the notable exception of a significant positive association for naturalistic exposure to educational TV. That is, children's vocabulary was not correlated with the overall amount of screen media exposure but correlated with exposure to educational media specifically. Given the near-zero average effect size for correlational studies, our findings do not support the *opportunity cost* hypothesis that screen media use has a net negative effect on vocabulary by displacing language-rich activities and social interactions. That is, even naturalistic exposure to TV/video content that was not explicitly designed to be educational was not negatively associated with vocabulary.

In balance, the results of this meta-analysis are consistent with a contingent *opportunity benefit* hypothesis in that screen media designed with educational intent may provide enriching opportunities for vocabulary learning even under conditions of naturalistic exposure. Other content may have minimal effects. Together with the finding that exposure dosage (i.e., number of exposures) did not predict the association between screen use and vocabulary, the results highlight the growing body of evidence that it is not the quantity but the quality of media that determines the extent and direction of impact on children's learning (for reviews, see Lauricella et al., 2016; Madigan et al., 2020).

Experimental versus naturalistic context

Our subsample analyses indicated a significant positive average effect for experimental studies ($r=.30$), but a non-significant average effect for correlational studies ($r=.07$). Before considering why these differences may have occurred, we note that the latter finding of a non-significant correlation is inconsistent with the significant *negative* association between screen use and children's general language skills found by Madigan et al. (2020) in their analysis of correlational studies ($r=-.14$). This discrepancy could reflect a number of differences between the two data sets. We only examined foreground media exposure; Madigan et al. included background media and found a negative association between this type of exposure and language skills ($r=-.19$). We found stronger positive effects for e-books and other interactive media than TV/video; they only examined time spent watching television, movies, and DVDs. We focused specifically on vocabulary; they included other aspects of language development (e.g., grammar, pragmatics) that may be more susceptible to opportunity costs of media use. Madigan et al. had a wider age range and a smaller average age compared to ours. Most notably, however, they also found a small positive correlation ($r=.14$) between children's exposure to educational media content and language skills, consistent with our findings for vocabulary ($r=.17$). As such, both meta-analyses suggest some value of educational media content for early language acquisition.

If correlational studies suggest overall minimal effects of naturalistic screen media exposure (or weak negative effects as observed by Madigan et al., 2020), why were there positive effect sizes for experimental exposure? One potential explanation is that experiments tended to expose children to media content created or edited by the researchers to teach specific words. In contrast, correlational studies assessed exposure to extant, professionally-produced media, often including general entertainment content. Unlike most professionally-produced content, researcher-created content often involved very simple visual presentations of an object accompanied by repetitions of a verbal label, minus any plot or story-line (e.g., Krcmar et al., 2007; Myers et al., 2017; Roseberry et al., 2014; Russo-Johnson et al., 2017). For instance, in Russo-Johnson et al. (2017), each novel word was repeated three times, “Look at the Dax! See the Dax? Isn't the Dax neat?” In Roseberry et al. (2014) there were 12 repetitions within a 1-min demonstration. It is possible young children perceive the educational intent of such content and invest more mental effort in learning from it (Field & Anderson, 1985; Salomon, 1984; Salomon & Leigh, 1984). If so, the relatively strong positive effects of such content suggest that professionally-produced content might have more positive effects on vocabulary if it was somewhat simpler with more repetition of explicit labeling.

Notably, while the effect sizes involving professionally-created media were significantly smaller compared to researcher-created media across all studies, this effect seems to be driven by the relatively smaller effect in correlational studies; media source was not a significant moderator for experimental studies. Moreover, we still found a positive average effect size for experimental studies involving professionally-produced media. Even among studies that used professionally-produced media, experimental studies generated significantly larger effects compared to correlational studies. Thus, media source alone cannot explain the difference between experimental and correlational studies. Among others, the type of vocabulary assessment may also contribute to this difference. In our dataset, experiments tend to assess learning of words depicted in the content rather than general vocabulary and, (unsurprisingly), there are stronger associations between media exposure and knowledge of words explicitly taught than between media exposure and general vocabulary (e.g., Smeets & Bus, 2015). Indeed, the relatively smaller effect for general vocabulary was only significant in the full sample of studies; in the experimental subsample, the effect for general vocabulary did not differ from zero or from the effect in correlational studies.

A third contributing factor to the different findings for experiments versus correlational studies may be the nature of effect sizes for these two methods. An experimental effect size represents the difference in vocabulary between children who were exposed to screen media and children who were not, while the correlational effect size denotes the change in vocabulary with an increase in screen media exposure. In this sense, while the experimental studies suggested that a particular screen media exposure was beneficial as compared to no exposure (i.e., a dichotomous variable), the findings presented in correlational studies assume a dose–response relation (i.e., a linear effect). Yet the association between naturalistic media use and general vocabulary development may not be linear. Indeed, one recent study reported a non-monotonic relation between weekly screen media time and vocabulary size gains in 6- and 8-year-olds, showing the largest vocabularies for children with intermediate screen time (Dore et al., 2020). In this way, treating screen exposure as a linear predictor may mask underlying associations with child outcomes. Thus, Dore et al.'s study, together with our findings, underscores the need for future research to consider nonlinear media effects.

A final contributing factor may be that experiments typically involve a reference group (i.e., no-exposure control) with constrained alternate activities, ensuring children cannot learn target vocabulary another way. By contrast, for correlational studies, the potential alternate activities in naturalistic environments could be just as (or more) beneficial for vocabulary learning (e.g., print book reading). As a result, children in correlational studies were perhaps learning some vocabulary from screen media, but due to a wide range of alternate learning activities, the net effect of the screen exposure was null. In this case, screen media use may represent a different (neither better nor worse) opportunity for learning.

Screen media platforms and features

In our sample, e-books were more beneficial than TV/video and games/apps for vocabulary acquisition. These findings echo the results of an earlier meta-analysis (Takacs et al., 2015), which found positive effects of e-books on preschool- and elementary-school-aged children's expressive vocabulary and story comprehension, (and even a small but significant advantage of e-books over print books). The current findings of stronger effects for e-books than for TV/video and games/apps may reflect the greater likelihood of interactive features and educational design to foster early literacy in e-books relative to other screen media. Relatedly, it is possible that children may more readily perceive the educational relevance of e-books than other screen media. Research has suggested that children's beliefs about the function of media could influence their attention to, interaction with, and learning from media, at least in the case of television (Field & Anderson, 1985; Salomon, 1984; Salomon & Leigh, 1984).

Unlike e-books, video chat did not have an overall positive effect on vocabulary learning. While professional organizations often list video chat as a uniquely beneficial form of screen media for young children (Chassiakos et al., 2016), we found that it was the only platform not significantly related to children's vocabulary in the full data set. An inspection of the data revealed that, while most effect sizes from individual studies were significantly positive, they had large variation due to factors including child age and vocabulary repetition. For instance, a study in which the program-specific words were repeated 12 times in a single demonstration for 30- to 42-months-olds generated a very large effect size ($r > .8$; Roseberry et al., 2014), whereas the effect size derived from a program-specific word repeated 4 times for 24-month-olds was close to zero (Troseth et al., 2018). Perhaps such large variation in a small sample of studies yielded a non-significant effect for video chat. Indeed, high variability across studies would explain why the media effect was not significant for video chat despite having the largest absolute correlation ($r = .46$) among all overall moderator categories.

Further, we found that interactivity moderated the average effect size in our sample of experimental studies. Given that we defined interactivity as having at least somewhat reciprocal interaction between the child and screen media, this finding complements the recent meta-analysis (Strouse & Samson, 2021) that found stronger effects for bidirectional (vs. one-way) live video on a variety of learning outcomes, including language learning, in 0- to 6-year-olds. Despite cases where interactive features are not designed appropriately or may disrupt learning (Furenes et al., 2021; Sheehan & Uttal, 2016), the current meta-analysis, by aggregating individual empirical studies, provides evidence of the effectiveness of interactivity, at least for learning program-specific vocabulary. As various authors have noted, the contingency afforded by interactive features may increase children's engagement, arousal, or selective attention and promote a sense of agency, which may increase engagement with and motivation for learning information (Kirkorian, 2018; Kuhl, 2007). Despite cases where interactive features are not designed appropriately and thereby disrupt learning (Choi & Kirkorian, 2016; Sheehan & Uttal, 2016), this meta-analysis, by aggregating individual empirical studies, provides evidence of the overall effectiveness of interactivity on vocabulary acquisition from screen media. Future empirical research should consider extending the facilitative role of interactivity media to naturalistic settings.

Child characteristics: Age and gender

We found that the relation between vocabulary and screen media exposure increased with children's age. That is, age was a significant continuous predictor of effect size across all studies and across experimental studies. Moreover, while the average effect size for screen exposure was significant among both younger and older children, it was significantly larger for those aged 36 months and older. This age group effect held in the subsample analysis for experimental studies and thus cannot be explained by uneven age distribution across research methods. Consistently, the meta-analysis of Madigan et al. (2020) based on children up to age 12 found a small positive association between age at time of exposure to screen media and vocabulary. Relatedly, Strouse and Samson (2021) found that young children's difficulty learning from videos (vs. real-life demonstrations) gradually diminished with age. The current findings suggest that, to the extent there are opportunity benefits associated with media use (e.g., word learning from media with educational intent), such benefits are likely to be smaller for infants and toddlers than for preschool-age children. This is consistent with established age-related increases in a wide range of cognitive skills, such as working memory, attention control, and symbolic thinking (Kirkorian, 2018; Strouse & Samson, 2021).

Child gender, categorized as the proportion of female participants, was the only other child characteristic with sufficient data for a moderator analysis. We found no evidence that the proportion of female participants moderated the size of the media effect in our dataset. Although Madigan et al. (2020) found that boys (vs. girls) showed greater gains in vocabulary when co-viewing with a caregiver, we were focusing solely on unaided media use, and did not observe gender differences for effects of such unaided exposure. The lack of gender effects may also partly reflect the fact that the majority of effect sizes in our sample came from experiments, where male and female participants tended to be evenly distributed. As such, we had relatively little variability in gender distribution.

Importantly, at the time of this writing, there was not a critical mass of studies or sufficient variability to enable moderator analyses for other individual or family-level differences. Prior work has suggested, for example, that educational media exposure was positively associated with early reading skills, even in the context of family stressors such as lack of economic resources (Vandewater & Bickham, 2004). Understanding more about the impact of screen media on early language in different family contexts would be an important goal for future research on screen media and vocabulary outcomes.

Receptive and expressive vocabulary

Finally, despite evidence that receptive vocabulary is easier to acquire compared to expressive vocabulary (Laufer & Goldstein, 2004), our analysis found larger effects of screen media exposure on expressive vocabulary. This finding was robust across research methods and media platforms. On the one hand, learning to express a word presumably needs an initial understanding of the word, which may explain the fact that receptive vocabulary precedes expressive vocabulary (Gershkoff-Stowe & Hahn, 2013; Stahl & Stahl, 2004). On the other hand, expressive vocabulary assessment is generally more complex and stringent than receptive vocabulary assessment. That is, word recall needs more precise association in memory between the word label and the corresponding meaning to be able to produce the correct word for the referent (Sénéchal, 1997; Stahl, 1999).

Moreover, prior research suggests that the acquisition of expressive vocabulary is more difficult compared to receptive vocabulary, particularly when children merely receive the information without active responding (Sénéchal & Cornell, 1993; Verhallen & Bus, 2010). Thus, it is especially noteworthy that our finding revealed a larger effect of screen media on expressive vocabulary than receptive vocabulary.

Research suggests a number of media techniques that should benefit expressive as well as receptive vocabulary, such as highlighting the connections between audio content and relevant visual details by showing moment-to-moment motion changes or zooming in on a particular portion of the scene (e.g., Calvert et al., 1982; Kirkorian & Anderson, 2018). Other research suggests that expressive vocabulary can be particularly enhanced by certain interactive techniques such as active prompts for the child to respond to questions (Sénéchal, 1997). Thus, the positive effects for both types of vocabulary might not be as surprising as initially appears. Further work could continue to probe whether expressive and receptive vocabulary acquisition might be differentially impacted by specific kinds of media experience.

Limitations and future directions

Several limitations should be acknowledged. First, our inclusion criteria were limited to four types of media platforms that are most common among the age group studied. Digital activities on other platforms, such as computer-based educational software, could have different impacts on child learning outcomes. Likewise, our exclusion criteria constrained this study to unaided screen media use and first-language vocabulary acquisition in typically developing children. Each criterion limited the generalizability of our findings, which needs further work to examine issues such as joint media engagement, second language learning, and neurodiversity. In addition, it is important to note that, with the growing use of digital stimuli in experimental studies, our search is likely to have overlooked many studies that delivered word-learning stimuli via screens but did not include one of our search terms in their title, abstract, or keywords. In this sense, the findings of this study should be interpreted in the context of literature with a focus on screen media.

Second, for some moderators, we had an unbalanced sample with very few effect sizes in certain categories (e.g., video chatting). Even though sensitivity checks were conducted to ensure the results were not driven by subcategories with the larger sample, inadequate variation in the small-sample categories presents a risk of violating the homogeneity of variance assumption and results in low power to detect moderator effects and disentangle potential confounds between moderators. Relatedly, the moderators we examined were not always balanced across each other. For example, the TV/video platform was more likely to be (although not always) non-interactive, whereas the other platforms were more likely to be interactive. To disentangle effects of interactivity and platform, future empirical research should manipulate interactivity within each platform, holding other factors (e.g., media content) constant. That would enable a future meta-analysis to examine the impact of interactivity within each platform, as has been done with eBooks (Takacs et al., 2015).

Additionally, being limited by the existing literature, the only individual differences we could examine systematically were child age and gender. Future work should consider age-related skills that drive developmental differences, as well as other individual and contextual factors (e.g., school attendance, household income, parental education) that would help to identify important moderators and speak to the generalizability of findings.

Third, our findings were subject to the quality of studies in the literature. To minimize potential confounds resulting from variation in research quality, we employed stringent criteria to exclude studies that presented inconsistent or ambiguous statistical reporting, did not have a valid comparison group (for experimental studies), or introduced systematic variability that could impact the effect sizes, such as joint media engagement with caregivers or peers. We also monitored the possible risk of bias by analyzing study quality moderators (e.g., publication status, inclusion of covariates). However, it remains possible that individual studies still suffered from other forms of bias, such as lack of power for particular analyses, or selective interpretation of data (Ferguson, 2015). Additionally, while most experimental studies allowed for precise measurement of media exposure, correlational studies generally relied on parent report of media exposure. Thus, the moderator of media content was coarsely coded as binary (educational or not), because insufficient detail was provided by the included studies to conduct a thorough content analysis and quality evaluation. For all these reasons, the quality of the extant literature tempers the implications of our findings.

Finally, regarding the moderating effect of age, the current analysis only focused on the age when the screen media exposure occurred, as defined in each study's method section. In most cases (with the exception of a few longitudinal studies), this was also the age when vocabulary assessment occurred. Given the small number of longitudinal studies, we did not examine the moderating effect of a delay between media exposure and vocabulary assessment. However, given the mixed findings regarding the relation between screen media use and early vocabulary from longitudinal research (e.g., Barr et al., 2020; Schmidt et al., 2009; Zimmerman & Christakis, 2007), future work should examine screen media and vocabulary longitudinally to examine developmental effects. Additionally, more longitudinal work and cross-lagged analyses are required to probe the causal mechanisms of screen media in naturalistic settings.

Summary

We found no evidence of a systematic negative association between screen media exposure and vocabulary, as would be expected if screen media invariably displace more educationally valuable activities for vocabulary learning. Rather, findings from this study demonstrate an overall positive relation between screen media exposure and vocabulary in children ages 6 years and under. This was true even for samples with an average age below 36 months, although screen media effects were larger for samples of older children. Moreover, the distinct effect sizes derived from experimental and correlational studies suggest the need for caution when extending results to naturalistic settings, particularly in the absence of adult involvement.

The overall positive relation between screen media and vocabulary increased with age but was significantly positive even in children below 36 months old. Nonetheless, more screen media exposure is not necessarily associated with better vocabulary outcomes: the quality of screen media matters. While educational TV/video, e-books, and interactive features were associated with higher vocabulary scores, we found no evidence that naturalistic exposure to TV/video in general (i.e., not specified as educational) had an overall positive effect on vocabulary. Our results also suggest positive effects for both receptive and expressive vocabulary with a larger effect on expressive vocabulary, at least when assessing whether children learn target words presented in media that are designed with explicit educational intent in the context of experimental research. Taken together, this study suggests the value of high quality, carefully designed content in supporting children's word learning, but points to the urgent need for empirical research that systematically investigates how to translate the promise of screen media from laboratory to naturalistic settings.

ACKNOWLEDGMENTS

The coding of this meta-analysis was supported in part by a Summertime Award for Research from the School of Human Ecology at the University of Wisconsin-Madison. The authors thank the many investigators who graciously responded our request for information and give special thanks to Gabrielle Strouse and Sarah C Kucker who generously provided unpublished data to support this project. Additional thanks to Aaron Qian Yao for his reliability coding assistance.

FUNDING INFORMATION

This study is partially supported by a Summertime Award for Research from the School of Human Ecology at the University of Wisconsin-Madison.

ORCID

Mengguo Jing https://orcid.org/0000-0001-6151-9613
Ting Ye https://orcid.org/0000-0001-6009-641X
Heather L. Kirkorian https://orcid.org/0000-0002-7990-7777
Marie-Louise Mares https://orcid.org/0000-0002-0811-3056

REFERENCES

Anderson, D. R., & Hanson, K. G. (2017). Screen media and parent-child interactions. In R. Barr, & D. N. Linebarger (Eds.), *Media exposure during infancy and early childhood: The effect of content and context on learning and development* (pp. 173–194). Springer.

Anderson, D. R., & Kirkorian, H. L. (2015). Media and cognitive development. In R.M. Lerner, L. S. Liben, & U. Mueller (Eds.), *Handbook of child psychology and developmental science* (7th ed., pp. 949–994). Wiley. https://doi.org/10.1002/9781118963418.childpsy222

Anderson, D. R., & Pempek, T. A. (2005). Television and very young children. *American Behavioral Scientist*, *48*, 505–522.

Apple. (2017). *iPad in education*. https://images. apple.com/education/docs/ipad-in-education-results.pdf

Aslanabadi, H., & Rasouli, G. (2013). The effect of games on improvement of Iranian EFL vocabulary knowledge in kindergartens. *International Review of Social Sciences and Humanities*, *6*, 186–195.

Barr, R. (2013). Memory constraints on infant learning from picture books, television, and touchscreens. *Child Development Perspectives*, *7*, 205–210. https://doi.org/10.1111/cdep.12041

Barr, R., Kirkorian, H., Radesky, J., Coyne, S., Nichols, D., Blanchfield, O., Rusnak, S., Stockdale, L., Ribner, A., Durnez, J., Epstein, M., Heimann, M., Koch, F.-S., Sundqvist, A., Birberg-Thornberg, U., Konrad, C., Slussareff, M., Bus, A., Bellagamba, F., … CAFE Consortium Key Investigators. (2020). Beyond screen time: A synergistic approach to a more comprehensive assessment of family media exposure during early childhood. *Frontiers in Psychology*, *11*, 1283.

Becker, B. J. (2005). Failsafe N or file-drawer number. In H. R. Rothsetin, A. J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis prevention, assessment and adjustments* (pp. 111–125). Wiley.

Becker, B. J., & Wu, M. J. (2007). The synthesis of regression slopes in meta-analysis. *Statistical Science*, *22*, 414–429.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Roth-stein, H. R. (2009). *Introduction to meta-analysis*. Wiley. https://doi.org/10.1002/9780470743386

Bowman, N. A. (2012). Effect sizes and statistical methods for meta-analysis in higher education. *Research in Higher Education*, *53*, 375–382.

Calvert, S. L., Huston, A. C., Watkins, B. A., & Wright, J. C. (1982). The relation between selective attention to television forms and children's comprehension of content. *Child Development*, *53*, 601–610.

Card, N. A. (2012). *Applied meta-analysis for social science research*. Guilford Press.

Chassiakos, Y. L. R., Radesky, J., Christakis, D., Moreno, M. A., & Cross, C. (2016). Children and adolescents and digital media. *Pediatrics*, *138*, e20162593.

Cheng, E. R., & Wilkinson, T. A. (2020). *Agonizing over screen time? Follow the three C's*. The New York Times.

Cheung, M. W. L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods*, *19*, 211–229. https://doi.org/10.1037/a0032968

Chinn, S. (2000). A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine*, *19*, 3127–3131.

Choi, K., & Kirkorian, H. L. (2016). Touch or watch to learn? Toddlers' object retrieval using contingent and noncontingent video. *Psychological Science*, *27*, 726–736.

Coburn, K. M., & Vevea, J. L. (2015). Publication bias as a function of study characteristics. *Psychological Methods*, *20*, 310–330.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge.

Cooper, H. (2009). *Research synthesis and meta-analysis: A step-by-step approach* (Vol. *2*). Sage.

De Jong, M. T., & Bus, A. G. (2002). Quality of book-reading matters for emergent readers: An experiment with the same book in a regular or electronic format. *Journal of Educational Psychology*, *94*, 145–155.

Dickinson, D. K., & Porche, M. V. (2011). Relation between language experiences in preschool classrooms and children's kindergarten and fourth-grade language and reading abilities. *Child Development*, *82*, 870–886.

Dore, R. A., Logan, J., Lin, T. J., Purtell, K. M., & Justice, L. (2020). Characteristics of children's media use and gains in language and literacy skills. *Frontiers in Psychology*, *11*, 2224.

Farkas, G., & Beron, K. (2004). The detailed age trajectory of oral vocabulary knowledge: Differences by class and race. *Social Science Research*, *33*, 464–497.

Ferguson, C. J. (2015). Do angry birds make for angry children? A meta-analysis of video game influences on children's and adolescents' aggression, mental health, prosocial behavior, and academic performance. *Perspectives on Psychological Science*, *10*, 646–666.

Field, D. E., & Anderson, D. R. (1985). Instruction and modality effects on children's television attention and comprehension. *Journal of Educational Psychology*, *77*, 91–100.

Furenes, M. I., Kucirkova, N., & Bus, A. G. (2021). A comparison of children's reading on paper versus screen: A meta-analysis. *Review of Educational Research*, *91*, 483–517.

Furuya-Kanamori, L., & Doi, S. A. (2016). Angry birds, angry children, and angry meta-analysts: A reanalysis. *Perspectives on Psychological Science*, *11*, 408–414.

Gershkoff-Stowe, L., & Hahn, E. R. (2013). Word comprehension and production asymmetries in children and adults. *Journal of Experimental Child Psychology*, *114*, 489–509.

Habeck, C. W., & Schultz, A. K. (2015). Community-level impacts of white-tailed deer on understorey plants in North American forests: A meta-analysis. *AoB Plants*, *7*, plv119. https://doi.org/10.1093/aobpla/plv119

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2021). *Doing meta-analysis with R: A hands-on guide*. Chapman & Hall/CRC Press.

Hassinger-Das, B., Brennan, S., Dore, R. A., Golinkoff, R. M., & Hirsh-Pasek, K. (2020). Children and screens. *Annual Review of Developmental Psychology*, *2*, 69–92.

Hirsh-Pasek, K., Zosh, J. M., Golinkoff, R. M., Gray, J. H., Robb, M. B., & Kaufman, J. (2015). Putting education in "educational" apps: Lessons form the science of learning. *Psychological Science in the Public Interest*, *16*, 3–34.

Hudon, T. M., Fennell, C. T., & Hoftyzer, M. (2013). Quality not quantity of television viewing is associated with bilingual toddlers' vocabulary scores. *Infant Behavior and Development*, *36*, 245–254.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*. Sage.

Kirkorian, H. L. (2018). When and how do interactive digital media help children connect what they see on and off the screen? *Child Development Perspectives*, *12*, 210–214.

Kirkorian, H. L., & Anderson, D. R. (2018). Effect of sequential video shot comprehensibility on attentional synchrony: A comparison of children and adults. *Proceedings of the National Academy of Sciences of the United States of America*, *115*, 9867–9874.

Kirkorian, H. L., Choi, K., & Pempek, T. A. (2016). Toddlers' word learning from contingent and noncontingent video on touchscreens. *Child Development*, *87*, 405–413.

Korat, O., & Shamir, A. (2012). Direct and indirect teaching: Using e-books for supporting vocabulary, word reading, and story comprehension for young children. *Journal of Educational Computing Research*, *46*, 135–152.

Krcmar, M. (2011). Word learning in very young children from infant-directed DVDs. *Journal of Communication*, *61*, 780–794.

Krcmar, M., Grela, B., & Lin, K. (2007). Can toddlers learn vocabulary from television? An experimental approach. *Media Psychology*, *10*, 41–63.

Kuhl, P. K. (2007). Is speech learning gated by the social brain? *Developmental Science*, *10*, v110–v120. https://doi.org/10.1111/j.1467-7687.2007.00572.x

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning*, *54*, 399–436.

Lauricella, A. R., Barr, R., & Calvert, S. L. (2016). Toddler learning from video: Effect of matched pedagogical cues. *Infant Behavior and Development*, *45*, 22–30.

Linebarger, D. L., & Vaala, S. E. (2010). Screen media and language development in infants and toddlers: An ecological perspective. *Developmental Review*, *30*, 176–202.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. SAGE Publications, Inc.

Madigan, S., McArthur, B. A., Anhorn, C., Eirich, R., & Christakis, D. A. (2020). Associations between screen use and child language skills: A systematic review and meta-analysis. *JAMA Pediatrics*, *174*, 665–675.

Mares, M. L., & Sivakumar, G. (2014). "Vámonos means go, but that's made up for the show": Reality confusions and learning from educational TV. *Developmental Psychology*, *50*, 2498–2511.

McDaniel, M. A., Rothstein, H. R., & Whetzel, D. L. (2006). Publication bias: A case study of four test ven- dors. *Personnel Psychology*, *59*, 927–953.

Mendelsohn, A., Brockmeyer, C. A., Dreyera, B. P., Fiermana, A., Berkule-Silbermana, S., & Tomopoulosa, S. (2010). Do verbal interactions with infants during electronic media exposure mitigate adverse impacts on their language development as toddlers? *Infant Child Development*, *19*, 577–593. https://doi.org/10.1002/icd.711

Morgan, P. L., Farkas, G., Hillemeier, M. M., Hammer, C. S., & Maczuga, S. (2015). 24-month-old children with larger oral vocabularies display greater academic and behavioral functioning at kindergarten entry. *Child Development*, *86*, 1351–1370.

Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods*, *11*, 364–386.

Myers, L. J., LeWitt, R. B., Gallo, R. E., & Maselli, N. M. (2017). Baby FaceTime: Can toddlers learn from online video chat? *Developmental Science*, *20*, e12430.

National Reading Panel Report. (2000). *Teaching young children to read*. The National Institute for Child Health and Development.

Neuman, S. B., & Kaefer, T. (2013). Enhancing the intensity of vocabulary instruction for preschoolers at risk: The effects of group size on word knowledge and conceptual development. *The Elementary School Journal*, *113*, 589–608.

Niessen, N., Strattman, K., & Scudder, R. (2011). The influence of three emergent literacy skills on the invented spellings of 4-year-olds. *Communication Disorders Quarterly*, *32*, 93–102.
Nikkelen, S. W. C., Valkenburg, P. M., Huizinga, M., & Bushman, B. J. (2014). Media use and ADHD-related behaviors in children and adolescents: A meta-analysis. *Developmental Psychology*, *50*, 2228–2241.
Pace, A., Alper, R., Burchinal, M. R., Golinkoff, R. M., & Hirsh-Pasek, K. (2019). Measuring success: Within and cross-domain predictors of academic and social trajectories in elementary school. *Early Childhood Research Quarterly*, *46*, 112–125.
Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press.
Paivio, A. (2007). *Mind and its evolution: A dual coding theoretical approach*. Erlbaum.
Paris, S. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly*, *40*, 184–202. https://doi.org/10.1598/RRQ.40.2.3
Peterson, R. A., & Brown, S. P. (2005). On the use of beta coefficients in meta-analysis. *Journal of Applied Psychology*, *90*, 175–181.
Rice, M. L., Huston, A. C., Truglio, R., & Wright, J. C. (1990). Words from "Sesame Street": Learning vocabulary while viewing. *Developmental Psychology*, *26*, 421–428.
Rice, M. L., & Woodsmall, L. (1988). Lessons from television: Children's word learning when viewing. *Child Development*, *59*, 420–429.
Rideout, V. (2017). *The common sense census: Media use by kids age zero to eight*. Common Sense Media.
Rideout, V., & Robb, M. (2020). *The commmon sense census: Media use by tweens and teens*. Common Sense Media.
Roseberry, S., Hirsh-Pasek, K., & Golinkoff, R. M. (2014). Skype me! Socially contingent interactions help toddlers learn language. *Child Development*, *85*, 956–970.
Roseberry, S., Hirsh-Pasek, K., Parish-Morris, J., & Golinkoff, R. M. (2009). Live action: Can young children learn verbs from video? *Child Development*, *80*, 1360–1375.
Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. Hedges (Eds.), *The handbook of research synthesis* (pp. 231–244). Russell Sage Foundation.
Rothstein, H. R., & Bushman, B. J. (2015). Methodological and reporting errors in meta-analytic reviews make other meta-analysts angry: A commentary on Ferguson (2015). *Perspectives on Psychological Science*, *10*, 677–679.
Russo-Johnson, C., Troseth, G., Duncan, C., & Mesghina, A. (2017). All tapped out: Touchscreen interactivity and young children's word learning. *Frontiers in Psychology*, *8*, 578.
Salomon, G. (1984). Television is "easy" and print is "tough:" the differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology*, *76*, 647–658. https://doi.org/10.1037/0022-0663.76.4.647
Salomon, G., & Leigh, T. (1984). Predispositions about learning from print and television. *Journal of Communication*, *34*, 119–135.
Sánchez-Meca, J., Marín-Martínez, F., & Chacón-Moscoso, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods*, *8*, 448–467.
Schmidt, M. E., Rich, M., Rifas-Shiman, S. L., Oken, E., & Taveras, E. M. (2009). Television viewing in infancy and child cognition at 3 years of age in a US cohort. *Pediatrics*, *123*, e370–e375.
Segal-Drori, O., Korat, O., Shamir, A., & Klein, P. S. (2010). Reading electronic and printed books with and without adult instruction: Effects on emergent reading. *Reading and Writing*, *23*, 913–930.
Segers, E., Nooijen, M., & de Moor, J. (2006). Computer vocabulary training in kindergarten children with special needs. *International Journal of Rehabilitation Research*, *29*, 343–345.
Segers, E., & Verhoeven, L. (2003). Effects of vocabulary training by computer in kindergarten. *Journal of Computer Assisted Learning*, *19*, 557–566.
Sénéchal, M. (1997). The differential effect of storybook reading on preschoolers' acquisition of expressive and receptive vocabulary. *Journal of Child Language*, *24*, 123–138.
Sénéchal, M., & Cornell, E. H. (1993). Vocabulary acquisition through shared reading experiences. *Reading Research Quarterly*, *28*, 360–374.
Sheehan, K. J., & Uttal, D. H. (2016). Children's learning from touch screens: A dual representation perspective. *Frontiers in Psychology*, *7*, 1220.
Silverman, R., Kim, Y. S., Hartranft, A., Nunn, S., & McNeish, D. (2017). Effects of a multimedia enhanced reading buddies program on kindergarten and grade 4 vocabulary and comprehension. *The Journal of Educational Research*, *110*, 391–404.
Silverman, R. D., Martin-Beltran, M., Peercy, M. M., Hartranft, A. M., McNeish, D. M., Artzi, L., & Nunn, S. (2017). Effects of a cross-age peer learning program on the vocabulary and comprehension of English learners and non-English learners in elementary school. *The Elementary School Journal*, *117*, 485–512.
Smeets, D. J., & Bus, A. G. (2015). The interactive animated e-book as a word learning device for kindergartners. *Applied PsychoLinguistics*, *36*, 899–920.
Snow, C. E., Burns, S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. National Academy Press.
Stahl, S. A. (1999). *Vocabulary development*. Brookline Books.
Stahl, S. A., & Stahl, K. A. (2004). Word wizards all! In J. F. Baumann, & E. J. Kameenui (Eds.), *Vocabulary instruction: Research to practice* (pp. 59–78). Guilford Press.
Strouse, G. A., O'Doherty, K., & Troseth, G. L. (2013). Effective coviewing: Preschoolers' learning from video after a dialogic questioning intervention. *Developmental Psychology*, *49*, 2368–2382.
Strouse, G. A., & Samson, J. E. (2021). Learning from video: A meta-analysis of the video deficit in children ages 0 to 6 years. *Child Development*, *92*, e20–e38.
Strouse, G. A., Troseth, G. L., O'Doherty, K. D., & Saylor, M. M. (2018). Co-viewing supports toddlers' word learning from contingent and noncontingent video. *Journal of Experimental Child Psychology*, *166*, 310–326.
Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using multivariate statistics* (Vol. *5*, pp. 481–498). Pearson.
Takacs, Z. K., Swart, E. K., & Bus, A. G. (2015). Benefits and pitfalls of multimedia and interactive features in technology-enhanced storybooks: A meta-analysis. *Review of Educational Research*, *85*, 698–739.
Troseth, G. L., Flores, I., & Stuckelman, Z. D. (2019). When representation becomes reality: Interactive digital media and symbolic development. *Advances in Child Development and Behavior*, *56*, 65–108.
Troseth, G. L., Strouse, G. A., Verdine, B. N., & Saylor, M. M. (2018). Let's chat: On-screen social responsiveness is not sufficient to support toddlers' word learning from video. *Frontiers in Psychology*, *9*, 2195.
Ulferts, H., Wolf, K. M., & Anders, Y. (2019). Impact of process quality in early childhood education and care on academic outcomes: Longitudinal meta-analysis. *Child Development*, *90*, 1474–1489.
Valkenburg, P. M. (2015). The limited informativeness of meta-analyses of media effects. *Perspectives on Psychological Science*, *10*, 680–682.
Van Assen, M. A., Van Aert, R. C., Nuijten, M. B., & Wicherts, J. M. (2014). Why publishing everything is more effective than selective publishing of statistically significant results. *PLoS One*, *9*, e84896.
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.
Vandewater, E. A., & Bickham, D. S. (2004). The impact of educational television on young children's reading in the context of family stress. *Journal of Applied Developmental Psychology*, *25*, 717–728.

Vandewater, E. A., Rideout, V. J., Wartella, E. A., Huang, X., Lee, J. H., & Shim, M. S. (2007). Digital childhood: Electronic media and technology use among infants, toddlers, and preschoolers. *Pediatrics*, *119*, e1006–e1015.

Vatalaro, A., Culp, A. M., Hahs-Vaughn, D. L., & Barnes, A. C. (2018). A quasi-experiment examining expressive and receptive vocabulary knowledge of preschool head start children using mobile media apps. *Early Childhood Education Journal*, *46*, 451–466.

Verhallen, M. J., & Bus, A. G. (2010). Low-income immigrant pupils learning vocabulary through digital picture storybooks. *Journal of Educational Psychology*, *102*(1), 54–61.

Verhallen, M. J., Bus, A. G., & de Jong, M. T. (2006). The promise of multimedia stories for kindergarten children at risk. *Journal of Educational Psychology*, *98*, 410–419.

Verhoef, R. E., Alsem, S. C., Verhulp, E. E., & De Castro, B. O. (2019). Hostile intent attribution and aggressive behavior in children revisited: A meta-analysis. *Child Development*, *90*, e525–e547.

Viechtbauer, W. (2007). Accounting for heterogeneity via random-effects models and moderator analyses in meta-analysis. *Zeitschrift für Psychologie/Journal of Psychology*, *215*, 104–121.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*, 1–48.

Viechtbauer, W. (2015). *Meta-analysis package for R*. Opgehaald van. http://cran.rproject.org/web/packages/metafor/metaforpdf

Viechtbauer, W., & Cheung, M. W. L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, *1*, 112–125.

Zimmerman, F. J., & Christakis, D. A. (2007). Associations between content types of early media exposure and sub- sequent attentional problems. *Pediatrics*, *120*, 986–992.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Jing, M., Ye, T., Kirkorian, H. L., & Mares, M.-L. (2023). Screen media exposure and young children's vocabulary learning and development: A meta-analysis. *Child Development*, *00*, 1–21. https://doi.org/10.1111/cdev.13927