



How educational are “educational” apps for young children? App store content analysis using the Four Pillars of Learning framework

Marisa Meyer, Jennifer M. Zosh, Caroline McLaren, Michael Robb, Harlan McCaffery, Roberta Michnick Golinkoff, Kathy Hirsh-Pasek & Jenny Radesky

To cite this article: Marisa Meyer, Jennifer M. Zosh, Caroline McLaren, Michael Robb, Harlan McCaffery, Roberta Michnick Golinkoff, Kathy Hirsh-Pasek & Jenny Radesky (2021) How educational are “educational” apps for young children? App store content analysis using the Four Pillars of Learning framework, *Journal of Children and Media*, 15:4, 526-548, DOI: [10.1080/17482798.2021.1882516](https://doi.org/10.1080/17482798.2021.1882516)

To link to this article: <https://doi.org/10.1080/17482798.2021.1882516>



View supplementary material [↗](#)



Published online: 23 Feb 2021.



Submit your article to this journal [↗](#)



Article views: 5616



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 64 View citing articles [↗](#)



How educational are “educational” apps for young children? App store content analysis using the Four Pillars of Learning framework

Marisa Meyer^a, Jennifer M. Zosh^b, Caroline McLaren^a, Michael Robb^c, Harlan McCaffery^a, Roberta Michnick Golinkoff^d, Kathy Hirsh-Pasek^e and Jenny Radesky^a

^aDepartment of Pediatrics, University of Michigan, Ann Arbor, USA; ^bDepartment of Human Development and Family Studies, Penn State University, Brandywine, Media, PA, USA; ^cCommon Sense Media, San Francisco, CA, USA; ^dSchool of Education, University of Delaware, Newark, DE, USA; ^eDepartment of Psychology, Temple University, Philadelphia, PA, USA

ABSTRACT

Experts have expressed concerns about the lack of evidence demonstrating that children’s “educational” applications (apps) have educational value. This study aimed to operationalize Hirsh-Pasek, Zosh, and colleagues’ Four Pillars of Learning into a reliable coding scheme (Pillar 1: Active Learning, Pillar 2: Engagement in the Learning Process, Pillar 3: Meaningful Learning, Pillar 4: Social Interaction), describe the educational quality of commercially available apps, and examine differences in educational quality between free and paid apps. We analyzed 100 children’s educational apps with the highest downloads from Google Play and Apple app stores, as well as 24 apps most frequently played by preschool-age children in a longitudinal cohort study. We developed a coding scheme in which each app earned a value of 0–3 for each Pillar, defining lower-quality apps as those scoring ≤ 4 , summed across the Four Pillars. Overall scores were low across all Pillars. Free apps had significantly lower Pillar 2 (Engagement in Learning Process) scores (t-test, $p < .0001$) and overall scores (t-test, $p < .0047$) when compared to paid apps, due to the presence of distracting enhancements. These results highlight the need for improved design of educational apps guided by developmental science.

ARTICLE HISTORY

Received 11 December 2019
Revised 22 January 2021
Accepted 25 January 2021


KEYWORDS

Educational Apps; early Childhood; digital Play; mobile Devices; Four Pillars of Learning; digital Media

Young children are now exposed to technology in a manner unlike any other generation. Recent estimates illustrate the universality of mobile device access, with approximately 98% of children ages eight and under living in a home with some type of mobile device (Rideout, 2017). Access to mobile devices is likewise commonplace in lower-income households, with 96% owning a mobile device (Rideout, 2017). Children ages eight and under spend an average of over two hours daily using screen media, with a significant portion of time spent on mobile devices (Rideout, 2017). While many children use mobile

CONTACT Jenny Radesky,  jradesky@med.umich.edu  Department of Pediatrics, University of Michigan, Address: 300 North Ingalls St, 1107, Ann Arbor, MI 48109 Ann Arbor, USA

This article has been republished with minor changes. These changes do not impact the academic content of the article.

 Supplemental data for this article can be accessed [here](#).

© 2021 Informa UK Limited, trading as Taylor & Francis Group

devices to stream videos through YouTube or Netflix, mobile applications (“apps”) that involve gaming or educational instruction are similarly popular.

Apps marketed as “educational” in commercial app stores advertise instruction on a wide range of fundamental academic skills, including counting, reading, and pattern recognition. Given the educational categorization of apps in prominent app stores, parents may expect that their children will develop the advertised early academic skills through using them. In interviews about their child’s media use, parents often expressed confidence in their child’s ability to learn from educational apps, possibly “better ... than from hands-on toys” (Radesky et al., 2016, p. 505); however, experts worry that these commercially available apps may not be designed to capitalize on how children learn, and may not achieve the advertised benefits (Hirsh-Pasek et al., 2015). Nonetheless, children’s educational apps are extremely popular, with some apps reporting more than 100,000,000 downloads.

Despite increasing use of mobile devices by children, there is limited research assessing the quality of design features and content presented in commercially available apps. A recent content analysis of advertising practices in the top-downloaded apps for children “5 and Under” found that many contained disruptive video advertisements or persuaded children to watch advertisements in exchange for rewards (Meyer et al., 2019, p. 33), both of which have the potential to distract young app users from an app’s underlying learning goals. A review of the top-downloaded, “literacy-focused” apps revealed that the majority of apps included “testing-based activities,” designed to only accept responses that could be scored as correct or incorrect; however, “open-ended designs” are similarly important for providing opportunities for exploratory learning without rigid evaluation (Vaala, Ly, & Levine, 2015, p. 5, 27). Furthermore, Callaghan and Reich (2018) conducted a content analysis of math and literacy apps marketed to preschoolers and determined that many evidence-based strategies for teaching preschool-age children were not utilized; they also found that app design features failed to match normative developmental capabilities for preschool-age users. Notably, there was a lack of scaffolded challenges in app activities, as well as frequent feedback mechanisms that may “undermine users’ intrinsic motivation” (Callaghan & Reich, 2018, p. 289).

Some studies suggest that preschool-age children can learn specific skills and concepts through interactive apps, such as problem-solving skills (e.g., Tower of Hanoi puzzles, Huber et al., 2014), foundational mathematics skills (Outhwaite, Faulder, Gulliford, & Pitchford, 2019), and vocabulary (Chiong & Shuler, 2010). However, the apps that have undergone empirical testing and evaluation – such as those from PBS KIDS and university research teams – may differ fundamentally from those promoted in app stores. There have been no comprehensive reviews of the educational quality of children’s apps marketed with a variety of educational objectives. Such a review is needed to help parents and early childhood practitioners be informed consumers, particularly since many children’s apps appear to have been designed as vehicles for advertising revenue (Meyer et al., 2019) or data trackers (Binns et al., 2018).

Additionally, it is unclear whether paid apps differ in quality from apps that can be downloaded for free. Higher-income parents (29%) report more frequently purchasing apps for their child compared to lower-income parents (14%) (Rideout, 2017). Moreover, lower-income parents have expressed more confidence in the educational value of these commercially available apps (Radesky et al., 2016). Access to high-quality educational

television programming is associated with cognitive and behavioral benefits for lower-income children (Christakis et al., 2013; Linebarger, Barr, Lapierre, & Piotrowski, 2014), and it is possible that free high-quality educational apps could provide similar benefits. However, if free educational apps are lower quality than those requiring payment, this may exacerbate the digital quality divide. Further research is needed to determine whether differences exist between the quality of free and paid apps (Callaghan & Reich, 2018).

The first aim of this study was to create a coding scheme based on the “Four Pillars,” introduced by Hirsh-Pasek, Zosh, and colleagues (2015) as a framework for evaluating app quality. The Four Pillars comprise interrelated aspects of design that can support or inhibit child learning. Pillar 1 (Active Learning) examines whether an app is “minds-on,” requiring thinking and intellectual effort, rather than simply engaging in cause-and-effect interactions with the app (p. 8). Pillar 2 (Engagement in the Learning Process) assesses the interactive enhancements of an app. It evaluates whether these features serve to engage the user in app activities versus distract the user, as occurs with interactive hotspots, excessive visual or sound effects, or disruptive advertisements. Pillar 3 (Meaningful Learning) examines whether the content of the app is meaningful to children’s everyday experiences and is taught in a manner that can be contextualized within existing knowledge. Pillar 4 (Social Interaction) assesses the degree to which children can interact meaningfully with characters through the app interface or with caregivers around the app. This pioneering framework has not been operationalized into a coding scheme to evaluate apps marketed as “educational” to children. Thus, we aimed to create a coding scheme rooted in these ideal standards for learning, which may help inform both the design of apps that truly facilitate optimal learning, as well as parents’ expectations for children’s learning from commercially available products.

Our second aim was to systematically evaluate the content and design presented in the top-downloaded educational apps in the Google Play and Apple app stores, and the most commonly played educational apps in an ongoing cohort study of preschool-age children. Our final aim was to evaluate differences in educational quality between free and paid apps. Given low-quality content found in prior analyses (Callaghan & Reich, 2018; Vaala et al., 2015), we hypothesized that many apps would receive low scores on each of the Four Pillars, particularly free apps, which contain more advertisements (Meyer et al., 2019).

Method

Overall study design

We conducted a content analysis of 124 apps marketed as educational to children from three sources: the top-downloaded educational apps from the Google Play store; the highest ranked educational apps from the Apple app store; and the apps most frequently used by preschool-age children collected in a longitudinal study assessing child development and mobile device use. The University of Michigan Medical School Institutional Review Board approved the aforementioned study, and deemed the content analysis exempt from human subjects review.

Developing the four pillars coding scheme

To create a reliable coding scheme based on the Four Pillars framework, we drafted an initial coding scheme based on the Hirsh-Pasek et al. (2015) article, and applied it to apps marketed to young children that we downloaded for a previous study (Meyer et al., 2019). The initial coding scheme was iteratively refined by the research team (MM, JR) by applying it to varied sets of children's apps and meeting to discuss coding uncertainties. For example, because the first coding set yielded consistently low scores, we then coded presumably higher-quality apps from Common Sense Media recommendation lists. Common Sense Media is an independent non-profit that reviews media for youth ages 2–18. Due to significant inter-app variability in each of the Pillars, we decided that the coding scheme would have ordinal codes (0–3) to represent a range of design qualities. Subsequent discussions with the Four Pillars authors (KHP, JZ, MR, RG) were conducted to refine the coding scheme and achieve consensus about difficult-to-code concepts, such as the “meaningfulness” of learning (Pillar 3). To improve external validity, collaborative meetings were held between co-authors from different disciplines (i.e., Pediatrics and Developmental Psychology), ensuring the same conceptual understanding of the coding scheme from multiple disciplinary perspectives. The final coding scheme is shown in [Appendix A](#). A final set of reliability testing with both Android and Apple apps demonstrated overall high reliability (Pillar 1 ICC = .869; Pillar 2 ICC = .947; Pillar 3 ICC = .657; Pillar 4 ICC = .758).

Identifying apps to code

We chose to review the top apps in the “education” sections of the Google Play and Apple app stores to replicate the experience of consumers seeking educational apps, thereby maximizing ecological validity. Although Hirsh-Pasek et al. (2015) assert that only apps that “support a learning goal” should be considered educational, they emphasize that app stores contain many apps marketed as educational that do not have clear learning goals (p. 4); we included these in our analytic sample.

Between November 26–29, 2018, we selected 50 apps from the Google Play store: 25 apps with the highest downloads (5,000,000–100,000,000+) from the “Top Free in Education” subcategory of the “Family” section, and 25 apps with the highest downloads (10,000–500,000+) from the “Top Paid in Education” subcategory. Paid apps ranged in price from \$0.99–\$4.99. To select apps, we ranked them based on number of downloads; if more than one app had the same range of downloads than needed to fill our app quota, we selected based on order displayed in the store. Apps were excluded from analysis if they were: an older version of another app on the list (e.g., *Farming Simulator 16* and *Farming Simulator 18*), platforms for multiple activities requiring credit card information (e.g., *ABCmouse*), or not intended for at-home use (e.g., *ClassDojo*).

We selected 50 apps from Apple's app store between March 22–2 April 2019, but used a slightly different selection strategy because this app store does not display the number of app downloads. Instead, we selected 25 apps from the “Top Free” subcategory of the “Education” section and 25 apps from the “Top Paid” subcategory, based on their ranking in each subcategory. Paid apps ranged in price from \$0.99–\$8.99. Similar to the Google Play selection process, apps that were classroom communication tools or not intended for

at-home use (e.g., *Google Classroom*), and large platforms containing multiple educational games without free trials (e.g., *IXL*) were excluded from this analysis. Additionally, 13 apps that were already selected from the Google Play store were not selected for the Apple list (e.g., *Peekaboo Barn*). Because 18 of the top 25 paid apps were by the same developer, Toca Boca, we decided to only code the top five of these apps to allow more variability in the sample; we wanted to describe a greater range of the design features that children are encountering when using popular educational apps. We chose not to compare educational quality scores between Google Play and Apple samples because of these differences in selection procedure.

App collection from longitudinal study

To describe the educational quality of apps actually played by young children – rather than those advertised or ranked highly in app stores – we selected a sample of the most frequently played apps in the first wave of data collection of the Preschooler Tablet Study (NICHD R21HD094051), which is described in greater detail elsewhere (Radesky et al., 2020). In this study, one week of mobile device sampling occurs through a passive sensing app for children with Android devices, and battery screenshot collection for Apple devices. Data from each source were analyzed to yield the 24 most commonly played Android and Apple educational apps. Through incorporating these apps, we aimed to further increase the ecological validity of our sample.

Data analysis

After reaching reliability on a subsample of apps, two researchers played each app for approximately 20 minutes and independently assigned a value of 0–3 for each Pillar, whereby 0 represents the lowest score and 3 represents the highest score. Researchers also recorded rationales for each score, describing how their observations from gameplay corresponded with a Pillar’s criteria to support their given score. Uncertainties in coding were resolved by consensus. We then calculated the means and frequencies of each Pillar score. We summed the Four Pillar scores for each app to calculate an overall educational quality score, and categorized apps as lower-quality if they earned a total score of ≤ 4 . For all Pillars, a score of 0 or 1 reflected design features that detracted from the app’s intended educational objective, whereas a score of 2 or 3 reflected design features that enhanced the app’s intended educational objective. Thus, a score of 4 was chosen as the cutoff in determining an app’s overall educational quality because it would indicate an app scored ≤ 1 on most of the Pillars. We then conducted bivariate analyses examining the frequency of Pillar scores by app payment status (free versus paid) using t-tests and Chi-square tests of association; we used Spearman’s correlation to determine inter-pillar correlations.

Results

Operationalizing the four pillars: coding scheme description

Pillar 1 (Active Learning - see Table [Appendix A](#)) is guided by whether the app activities involve the child generating responses and ideas, versus simply reacting to on-screen

stimuli. This Pillar assesses whether the child will be cognitively challenged by the app activities, or whether the app includes only simple cause-and-effect design. A score of 0 indicates that activities could be completed with little mental effort, such as *Learning Letters Puppy*, which involves tapping the screen to see and hear a narrated alphabet or number line; design constraints include inability to go backwards or have letters repeated [Supplemental Material 1]. Other examples of a “0” code include bubble popping games, which do not require more mental effort than a simple reaction to engage with the app activities. Conversely, apps would score a 3 if the activities require the player to generate responses with incrementally more challenge, thereby eliciting “minds-on” learning (Hirsh-Pasek et al., 2015, p. 8). *Lego DUPLO Town* requires the player to independently generate responses to design aspects of their virtual buildings, rather than following instructions. Flexible challenges are provided. For example, if a building piece is placed outside of bounds, it will not snap into the closest spot; instead, it will return to the pieces bank, and the player must figure out why their construction did not work as intended [Supplemental Material 2]. An app would score a 1 if the activities presented a learning objective in a constrained format (e.g., completing a simple puzzle), whereas an app would score a 2 if the learning objectives were presented with a greater degree of flexibility, yet still included programmed instructions that guided overall gameplay.

Pillar 2 (Engagement in the Learning Process - see Table Appendix A) questions whether the interactive and gamified qualities of the app support engagement with the app’s learning objectives, or potentially distract the user. This Pillar assesses the prevalence, quantity, and distractibility of advertising, hotspot enhancements, and rewards that could themselves become the focus of gameplay. An app would earn a score of 0 if there were four or more types of disruptive advertisements, video advertisements were embedded in gameplay, a pop-up advertisement interrupted mid-gameplay, or the player was rewarded with tokens for every action. For example, in *Baby Panda Home Safety*, pop-up advertisements frequently interrupted the user’s actions mid-gameplay. We chose to assign low Pillar 2 (Engagement in the Learning Process) scores to apps with highly prevalent advertising because extant research maintains that too many distractions in educational material are negatively associated with children’s learning outcomes, likely due to overwhelming a child’s cognitive load and inhibiting their ability to process the instructed concepts (Barr, Shuck, Salerno, Atkinson, & Linebarger, 2010; Tare, Chiong, Ganea, & Deloache, 2010).

Pillar 2 (Engagement in the Learning Process) scores were further determined by the quantity of relevant feedback provided to the user, defined as corrective or formative responses contingent to the child’s actions (e.g., explaining why their answer was incorrect; see Falloon, 2013). Apps scored lower if they provided either nonspecific feedback (e.g., sound and visual effects) to user actions, or displayed excessive rewards (e.g., stars covering the screen) not relevant to the learning goals. For example, in *Sweet Baby Girl Daycare*, the player earns coins and stars after each step in a simple cooking activity; the player observes the tokens fly across the screen to their progress trackers and hears a “cha-ching” noise [Supplemental Material 3]. Conversely, an app would score a 3 if no advertising was present, or if enhancements were designed in a manner that supported the learning goals. For example, in *Toca Lab: Elements*, the characters make facial expressions and noises in reaction to the lab equipment used to transform them into new elements, thereby embellishing the open-ended play experience of the app. An app

would score a 1 if one-to-three types of advertisements were present that interrupted the child's activities, or if excessive rewards were present. For example, *Monkey Preschool Fix-It* scored a 1 for Pillar 2 (Engagement in the Learning Process) because excessive, irrelevant enhancements occurred after every action. Specifically, the monkey would smile, jump, and/or squeal after correct answers, and shake its head, wave its hands, and/or make a noise after an incorrect response [Supplemental Material 4]. Additionally, the player would hear a narrated affirmation after correctly completing each activity, frequently followed by a sparkly, rainbow trail around correct responses, or earning virtual toys after completing activities. In contrast, *Bedtime Math* earned a score of 3 for Pillar 2 (Engagement in the Learning Process) because relevant feedback was provided after the player answered a level's question; the correct answer slides across the screen while a slide-whistle noise plays, but no other hotspots or rewards are provided [Supplemental Material 5].

Pillar 3 (Meaningful Learning - see Table Appendix A) assesses whether the presentation of the educational curriculum promotes transfer of meaningful information or instructed skills to the child's life context. An app would score 0 if the activities had no clear learning objective or were not relevant to the child's life experiences. For example, *Baby Phone* prompted simple tapping on number and animal buttons to hear piano sounds or nursery songs [Supplemental Material 6]. There was no connection between the app's activities, visual effects, and sound effects to corresponding concepts in real life. Notably, if the player tapped on the call button, the screen would switch to a cartoon picture that moves side-to-side while a song plays, rather than demonstrating how a phone is used. For an app to score a 3, the activities should be based on learning experiences that would be easy for the child to replicate in real life, with gameplay strongly connected and relevant to the child's life. *Little Panda Travel Safety* teaches strategies for staying safe while away from home, demonstrating and narrating how the player can use the instructed skills (i.e., "whenever you go out, you must stay within your parent's sight"). The activities in this app are based on experiences that are not only relevant to a preschooler's life, but are presented in such a way that prompts the child to use them in real-life experiences. Relevance to children's lives had been interpreted as content that would be taught in preschool programs or would be beneficial for a preschool-age child to know, such as foundational skills like numeracy and phonics. Moreover, opportunities for pretend play, creative expression, and strategic problem-solving were likewise interpreted as relevant to a child's life. An app would score a 1 if the activities involved a meaningful concept for children (e.g., color recognition), but only provided a constrained set of activities to learn these concepts. An example of these gameplay constraints includes repeatedly sliding colored blocks into holes outlined in the corresponding colors, while hearing the color names narrated. Apps would score a 2 if they provide multiple ways to contextualize a learning concept. For example, an app may show how a letter fits into several different words and defines those words through animations, but does not teach or prompt the child to apply that knowledge to real life.

Pillar 4 (Social Interaction - see Table Appendix A) evaluates the level of social interaction promoted in gameplay. Social interactions may occur through co-playing with another person, or through development of parasocial relationships with characters. Parasocial relationships are one-sided, emotionally charged connections an individual forms with a media character which have been found to improve children's learning from

media (Bond & Calvert, 2014). If an app included no prompts to co-play and had no characters with which children could develop parasocial relationships, it scored 0 (e.g., *Drawing for Kids and Toddlers*). An app earned a 3 if it provided explicit opportunities for social engagement around the app and presented ideas for how to transfer the instructed concepts to the real world with another person (e.g., “Pinkamazing Family Game” of *PBS KIDS Games*). If an app’s structure provided opportunities for development of weak parasocial relationships with app characters or had excessive, gamified features distracting from socially interactive features, it scored a 1. The characters of *Dentist Games for Kids* smiled when the player finished cleaning their teeth, but did not comment or act in a manner suggesting that the character understood the player’s actions [Supplemental Material 7]. An app earned a 2 if it provided opportunities for social interaction with another person, such as through multi-touch input in an activity. An app also earned a 2 if it provided opportunities for development of high-quality parasocial relationships with app characters, such as when a character responds to the child’s actions in ways that show they understand the child. For example, in *XtraMath* the teacher reacted in a way that made it seem as though he understood the player’s actions through how he discussed the player’s quiz results. See Appendix A for the summarized coding scheme.

Evaluation of the content and design of apps: descriptive statistics

Scores for Pillars 1–3 ranged from 0 to 3, spanning the entire quality spectrum assessed by the coding scheme. Pillar 4 (Social Interaction) scores only ranged from 0 to 2, for both free and paid apps (Figure 1).

A score of 1 was the most frequent score earned by apps for Pillar 1 (Active Learning), regardless of free versus paid distinction. Eighty-one of 124 apps (65%) scored a 1,

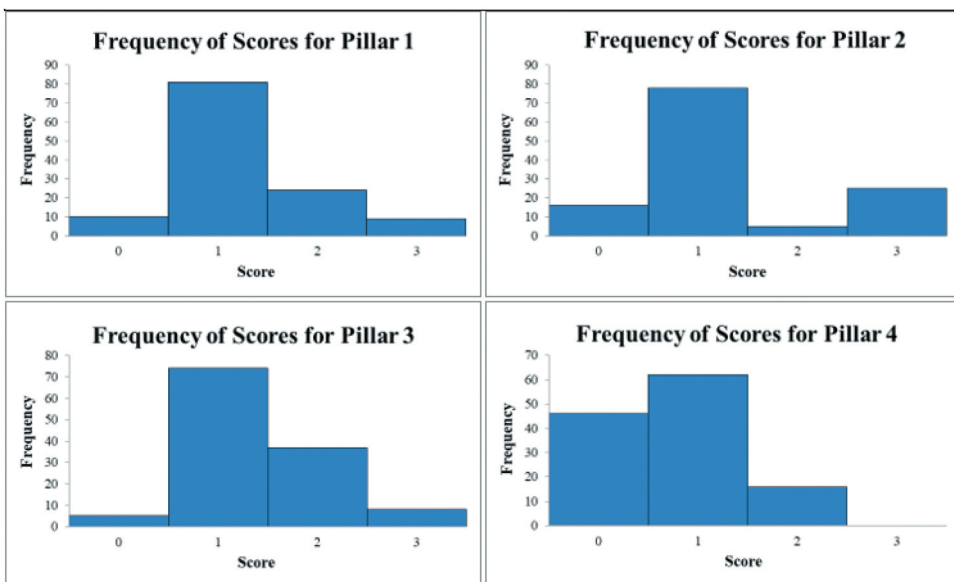


Figure 1. Frequency of codes for the four pillars.

indicating widespread structural constraints and a low degree of cognitive effort promoted by the apps in our sample. Ten additional apps (8%) scored a 0, leading to 91 apps (73%) demonstrating lower-quality design as assessed by Pillar 1 (Active Learning) criteria. Likewise, a score of 1 was the most frequent score earned for Pillar 2 (Engagement in the Learning Process), with 78 apps (63%) earning this score. Sixteen apps (13%) also earned a 0, resulting in 94 apps (76%) presenting lower-quality interactive design; this suggests disruptive advertising and excessive feedback mechanisms were prominent throughout sampled apps, with the potential to distract users from gameplay. A score of 1 was the most frequent score earned for Pillar 3 (Meaningful Learning; 74 apps, 60%), with an additional five apps (4%) scoring a 0. Seventy-nine apps (64%) earning lower-quality scores for Pillar 3 (Meaningful Learning) suggests that the designs of sampled apps did not promote meaningful learning in a digital setting, nor support users in transferring learned skills to another setting. Sixty-two apps (50%) earned a score of 1 for Pillar 4 (Social Interaction), followed by 46 apps (37%) scoring a 0; the frequency of these lower-quality scores suggests that 87% of apps in our sample failed to promote opportunities for social interaction. Figure 1 illustrates the frequency of scores for each Pillar. Of 124 apps sampled, 72 (58%) earned a total score of ≤ 4 , thereby considered lower-quality.

Quality differences based on payment

Table 1 provides means and standard deviations for each Pillar based on app payment status. Low mean scores were observed across all Pillars. Pillar 2 (Engagement in the Learning Process) scores were significantly lower in free apps compared to paid apps ($t(122) = 3.97, p < .0001$), as were overall scores ($t(122) = 2.88, p < .0047$). Additionally, more free apps (47 apps, 64%) were considered lower-quality than paid apps (25 apps, 50%), although differences were not statistically significant ($\chi^2(1, N = 124) = 2.24, p = .14$).

Inter-pillar correlations

Pillar 1 (Active Learning), Pillar 2 (Engagement in the Learning Process), and Pillar 3 (Meaningful Learning) were significantly associated with each other; however, Pillar 4 (Social Interaction) was not significantly associated with the other Pillars. Specifically, Pillar 1 (Active Learning) was moderately correlated with Pillar 2 (Engagement in the Learning Process) ($r_s = .36, p < .0001, N = 124$) and Pillar 3 (Meaningful Learning) ($r_s = .55, p < .0001, N = 124$). Pillar 2 (Engagement in the Learning Process) was significantly correlated with Pillar 3 (Meaningful Learning), albeit to a smaller degree ($r_s = .26, p = .0035, N = 124$).

Table 1. Mean 6pillar scores by payment status.

Pillar	Free Apps Mean (SD)	Paid Apps Mean (SD)
Pillar 1 – Active Learning	1.18 (0.63)	1.38 (0.81)
Pillar 2 – Engagement in the Learning Process	1.05 (0.77)*	1.70 (1.04)*
Pillar 3 – Meaningful Learning	1.35 (0.71)	1.44 (0.61)
Pillar 4 – Social Interaction	0.73 (0.67)	0.80 (0.67)
Total Score	4.31 (1.69)**	5.32 (2.21)**

* $p < .0001$, ** $p < .0047$

Discussion

In this study, our primary objective was to assess the educational quality of commercially available children's apps based on the Four Pillars framework (Hirsh-Pasek et al., 2015). To accomplish this goal, we created a detailed coding scheme based on established principles for facilitating learning in early childhood.

Overall, we found that most of the top-downloaded children's educational apps from Google Play, Apple, and our sampled cohort's selection scored in the lower range of Four Pillars principles, regardless of payment status. However, free apps contained more distracting visual and sound effects, disruptive advertising, and irrelevant rewards than paid apps, as illustrated by lower Pillar 2 (Engagement in the Learning Process) scores; these features may distract from other educational objectives. Furthermore, even in paid apps, which parents may assume are higher-quality, 50% of our sample scored in the lower-quality range (≤ 4). Seven apps (6%) earned a total score greater than 8, suggestive of a higher-quality educational experience in the app. The highest total score earned was a 10, of which two paid apps (*Daniel Tiger's Stop & Go Potty* and *My Food–Nutrition for Kids*) earned.

This study contributes to the limited research on children's educational apps. To the best of our knowledge, this is the first study assessing the quality of content and design features in the top-downloaded educational apps from Google Play and Apple, as well as apps actually played by preschool-age children. It is likewise one of few studies examining quality differences between free and paid educational apps (Callaghan & Reich, 2018). Prior research has primarily focused on the curricular content of apps for only a single domain, such as literacy or math skills. Nonetheless, in accordance with prior content analyses, we found high levels of advertising (Meyer et al., 2019) and distracting, tangential animations (Vaala et al., 2015), as evidenced by lower Pillar 2 (Engagement in the Learning Process) scores. Moreover, lower Pillar 4 (Social Interaction) scores highlight the few embedded opportunities for in-person or mediated social interaction in popular educational apps; Vaala et al. (2015) revealed a comparable dearth of apps "designed to promote joint media engagement" (p. 6). Research examining parent–child interactions around tablets has similarly highlighted limitations in parent–child reciprocity when apps have more enhancements or fast-paced design (Hiniker, Lee, Kientz, & Radesky, 2018; Munzer, Miller, Weeks, Kaciroti, & Radesky, 2019).

Inter-pillar correlations were small to moderate among Pillar 1 (Active Learning), Pillar 2 (Engagement in the Learning Process), and Pillar 3 (Meaningful Learning); however, Pillar 4 (Social Interaction) was not significantly associated with the other Pillars. These findings suggest that high-quality, well-designed apps are more likely to be well designed along multiple design facets. For example, an app that promoted greater "minds-on" thinking (Hirsh-Pasek et al., 2015, p. 8) also likely had more formative feedback and fewer disruptive advertisements, scoring higher on both Pillar 1 (Active Learning) and Pillar 2 (Engagement in the Learning Process). Lack of significant correlations between Pillar 4 (Social Interaction) and the other Pillars likely reflects the rarity of structured opportunities for social interaction in currently available commercial apps.

Challenges in operationalizing and implementing the coding scheme

While developing the Four Pillars coding scheme, we encountered challenges with applying this theoretical framework to actual app experiences and achieving consistent observations between different devices; these challenges hold relevance for other research teams considering app content analysis.

Pillar 2 (engagement in the learning process) – changing advertisement patterns

Time of day, date of gameplay, and duration of gameplay were not held constant when coding apps; thus variability emerged in the quantity and types of advertising each researcher observed. Variability in advertising practices was primarily observed among the disruptive advertising criteria that characterized the lower-quality scores of 0 and 1, such as the presence of pop-up video advertisements and rating prompts. For example, one researcher might experience a pop-up video advertisement after exiting an activity, while another researcher experienced a pop-up advertisement interrupting gameplay because they remained in an activity longer than the first researcher. These inconsistencies may be due to differences in how advertising networks populate advertisements. However, given the high reliability for Pillar 2 (Engagement in the Learning Process), we suspect that variability in advertising did not significantly distort the consistency of sampled gameplay experiences.

Pillar 2 (engagement in the learning process) – emphasis on distractions versus feedback quality

We intended for Pillar 2 (Engagement in the Learning Process) criteria to evaluate the extent to which interactive features support or distract the user from engaging with an app's learning objectives. Although prior studies have detailed the differences between quality and contingency of digital scaffolds (Callaghan & Reich, 2018), we found that the sheer quantity of advertising and enhancements in sampled apps precluded a close examination of digital scaffold quality. We judged that distracting advertisements, tokens, virtual toys, and other irrelevant design approaches would invoke children's extraneous processing, in which these tangential design features tax working memory capacity and strain a child's cognitive capacity for learning new information (Mayer, 2014). Moreover, prior research emphasizes the importance of the visual environment for learning and children's abilities to both selectively attend to and sustain attention on tasks. For example, kindergarteners in classrooms with many colorful decorations have greater difficulty focusing on instructed tasks than those in classrooms without decorations (Fisher, Godwin, & Seltman, 2014), and preschoolers perform worse on tasks when distractions are continuously presented (Kannass & Colombo, 2007); the same may apply to distracting digital environments.

We also considered the extent to which excessive enhancements and rewards for user actions might undermine an app's learning objectives. Although feedback coupled with extrinsic rewards are helpful when relevant to the desired behavior (e.g., "natural rewards"), irrelevant rewards (e.g., virtual toys for completing math equations) and excessive enhancements (e.g., multiple forms of praise immediately after any user action) have been hypothesized to accustom children to extrinsic motivation, rather than develop intrinsic motivation to complete tasks (Hirsh-Pasek et al., 2015). More research is needed to understand the balance of digital enhancements and rewards in supporting versus distracting from child learning.

Pillar 3 – what constitutes “meaningful learning?”

Throughout the process of developing the coding scheme, defining the concepts involved in Pillar 3 (Meaningful Learning) posed the greatest challenge. For example, are simple concepts (e.g., counting) more meaningful to app users at earlier developmental stages than to more mature app users? Are apps that provide “pretend play experiences,” such as a hair salon or hospital, only meaningful if they allow significant autonomy and imagination for the child?

The value derived from rote learning differs by child and academic subject, which was difficult to capture in this coding scheme. Rote learning can lead to memorization of foundational concepts and facts, but deeper conceptual understanding of instructed material improves retention of concepts and facilitates transfer to other contexts or tasks (Committee on Developments in the Science of Learning, 2000). Thus, our coding scheme emphasized the extent to which an app provides these meaningful learning opportunities, including instruction of concepts in multiple contexts rather than via a single example (Committee on Developments in the Science of Learning, 2000), or encouraging a child to apply new information to their everyday experiences (Hirsh-Pasek et al., 2015; Vosniadou, 2001). For instance, an app intended to teach the alphabet would promote deeper conceptual understanding of a given letter through repeating the letter sound, tracing the letter, and showcasing pictures and the corresponding spelling of words that start with the letter.

Pillar 4 (social interaction) – assessing the quality of parasocial relationships

When developing measures of social interactivity, we not only considered the possibility of social engagement with other players in-person around the device or mediated by the device, but also considered the possibility of social engagement with app characters via parasocial relationships. We struggled to judge the degrees of quality for parasocial relationships, as the mere presence of an app character standing adjacent to an activity (e.g., a cartoon lion in *ABC Kids–Tracing and Phonics* that stands next to the to-be-traced letter and looks around while waiting for the player to trace the letter) differs from an app character that “talks” to the player through programmed responses to the player’s actions. Additionally, prior relationships with branded characters, such as Elmo and the Super Why! characters, may influence the player’s engagement with the character, compared to a novel character created for a specific app. We decided to include both degree of contingent interactivity and familiarity of characters when assessing the quality of parasocial relationships in our coding scheme. For example, *Teenage Mutant Ninja Turtles: Half-Shell Heroes* includes popular media characters with whom preschool-age children are likely familiar, but they provide only generic, affirmative feedback to the player’s actions, yielding a lower Pillar 4 score. Conversely, in *Daniel Tiger’s Stop & Go Potty*, the well-known character Daniel Tiger expresses greater understanding of the player’s actions through tailored feedback, leading to development of a stronger parasocial relationship.

Finding a balance in the level of detail involved in coding

A primary challenge in developing and using the present coding scheme involved its high level of detail. Our coding scheme’s range of 0–3 was intended to prevent oversight of important details that would distinguish higher-quality apps from lower-quality apps, but the nuances in the coding scheme led to more time needed to achieve inter-rater reliability. Compared to television research – in which educational content coding is

performed on a limited number of children's programs viewed by multiple participants – the sheer number of different children's apps is much higher; coding hundreds of apps would not be feasible when the coding process takes 20 minutes per app.

Future research teams may consider using binary 0/1 coding for simpler implementation. Moreover, a dichotomous coding scheme may be easier to reliably train multiple labs to use. In posthoc analyses, we condensed the four-tier coding scheme into a binary coding scheme (scores of 0 and 1 collapsed to a score of 0 to indicate lower-quality, and scores of 2 and 3 collapsed to a score of 1 to indicate higher-quality). With this dichotomous coding scheme, results were similar in terms of overall app quality (scores of 0 occurred in 73% of Pillar 1, 76% of Pillar 2, 64% of Pillar 3, and 87% of Pillar 4) and comparisons between free and paid apps.

Implications of quality differences based on payment

App quality differences based on payment status were observed for both Pillar 2 (Engagement in the Learning Process) and overall scores, whereby free apps earned lower scores than paid apps. As aforementioned, the driving factor of lower Pillar 2 (Engagement in the Learning Process) scores was the high prevalence of disruptive advertisements and frequent, reward-driven feedback; these design features may inhibit a child's ability to attend to and comprehend an app's educational content. If free apps are more frequently played by children of lower-socioeconomic status (SES), the lower-quality design we documented among free apps may serve to widen the present digital divide. Moreover, if children of lower-SES are using free educational apps for supplemental or preparatory learning experiences for grade school, the intended gains in early academic skills and subject matter may not be realized, given the disconnect between the design of popular educational apps and the features most conducive to children's learning.

Importantly, some apps are free because the developer incorporates in-app purchases and advertisements, whereas other apps are free because they are created by non-profit organizations. This distinction in business models may obscure quality differences for Pillar 2 (Engagement in the Learning Process) based solely on payment status. Within our sample, PBS KIDS apps were the only free apps developed by non-profits. In post hoc analyses, PBS KIDS apps ($n = 3$) received higher scores than free apps ($n = 72$) for almost all Pillars (Pillar 1: PBS $M = 1.00$, Free $M = 1.19$; Pillar 2: PBS $M = 1.67$, Free $M = 1.06$; Pillar 3: PBS $M = 2.33$, Free $M = 1.33$; Pillar 4: PBS $M = 2.00$, Free $M = 0.69$). However, the small sample size precludes inferential statistical analyses.

Additional implications for research

Despite the time required to code an app using a theory-based coding scheme, the evaluation of the quality of children's educational apps is necessary to establish greater understanding of the quality of interactive and mobile media played by children. The developmental benefits associated with high-quality children's media content are well established; however, research has largely concentrated on the quality of educational television programming (Linebarger, Brey, Fenstermacher, & Barr, 2017). Content analyses have described in detail the strengths of various instructional methods for character

interactivity, as well as how structural elements like pace and context familiarity affect a child's ease in processing the educational material presented. High-quality educational television programs have been associated with increased "school-readiness, ... problem solving ... and literacy" for viewers (Linebarger et al., 2017, p. 98). Yet children now spend a greater share of their average media consumption using mobile devices (Rideout, 2017); when operating as a platform for app use, the structure and interactive features of mobile devices differ considerably from traditional television programs. However, little research exists detailing how educational apps compare to educational television programming, and whether the described positive developmental outcomes exist for educational apps. Moreover, this evaluation matters because time spent playing lower-quality educational apps could be displacing more developmentally beneficial activities, such as reading, physical activity, or unstructured social play (AAP Council on Communications and Media, 2016).

Additionally, the Four Pillars approach allows for evaluation of different aspects of app design that may have varying implications for child development and gameplay experience. For example, Pillar 1 (Active Learning) may relate more to open-ended problem-solving, while Pillar 2 (Engagement in the Learning Process) may matter more for child distractibility; Pillar 4 (Social Interaction) may have implications for social reciprocity. Importantly, an app does not necessarily need to earn top scores for every Pillar to be considered a worthwhile educational app. Future research should examine how Four Pillars scores associate with children's acquisition of in-app educational content, as well how each of the Pillars independently correspond with the aforementioned learning outcomes.

Implications for app designers, app stores, and consumers

These results highlight the need for improved design of educational apps by developers, as well as modifications to how educational apps are displayed and promoted in app stores. We encourage app designers to develop apps rooted in the ways children learn most effectively. Based on evidence from the Science of Learning, the Four Pillars framework aims to provide exactly that type of guidance (Hirsh-Pasek et al., 2015); however, work remains to translate this evidence in a way that is easily accessible and actionable by app designers. We recommend that app designers and app stores work with child development experts to create evidence-based ratings of apps, so that higher-quality products with fewer distracting enhancements can be found by parents.

These findings also suggest that parents may benefit from understanding how to evaluate the quality of educational apps before downloading them. The Zero to Three Screen Sense E-AIMS guidelines, rooted in the Four Pillars, provide instructions for parents on evaluating whether apps meet the Four Pillars criteria; this may help them avoid installing lower-quality apps (Zero to Three, 2018). Additionally, parents can be encouraged to install apps for their children, rather than letting children download apps on their own, or require a password before download. They may also seek app reviews from trusted sources like Common Sense Media when choosing content for their child. Likewise, parents can judge the quality of apps by co-playing with children, as well as help them build digital literacy and transfer their knowledge from apps to the three-dimensional world more readily (Barr, 2013). Lastly, these results highlight that most apps

cannot recreate more traditional forms of child-led, hands-on, unstructured play and their benefits, which should still be encouraged (Yogman, Garner, Hutchinson, Hirsh-Pasek, & Golinkoff, 2018).

Study limitations

Given the ever-changing nature of media, our study only describes a sample of educational apps from a specific point in time. Apps are frequently updated by developers; thus, the design of activities, curriculum, and gamified mechanisms presented in apps during the time-frame in which we coded them may be different if accessed at another time. Moreover, our coding scheme used a scale developed based on the range of observed design features in this sample; our cut-offs for different scores could not be informed by evidence about child learning, or compared to other educational experiences. Thus, our scores should not be interpreted as absolute measures of the educational quality of sampled apps. However, we were able to assess to what extent popular commercially available apps aligned with Four Pillars principles. This coding scheme represents a preliminary step in the process of developing reliable and valid measures for evaluating children's educational apps.

Additionally, our coding scheme did not account for the age of an app's target audience. Future investigators should consider incorporating this criterion in subsequent appraisal tools, as possible applications of the Four Pillars likely differ based on a user's age. For example, the structure and design features necessary to promote active learning (Pillar 1) for preschool-age children may be insufficient to promote active learning for school-age students.

Conclusion

Our study results highlight a need for improvement to the quality and design of interactive features in educational apps marketed to young children. Overall, 58% of apps studied showed lower-quality design when assessed against the Four Pillars of Learning principles. Although smartphones and tablets have been widely available only in the last decade, it is important for the industry to join with cognitive and developmental experts on children's learning to produce apps that are more consonant with how children learn. Until then, it is important for parents to know that not all apps are equally "educational."

Funding

This work was supported by the National Institute of Child Health and Human Development [NICHD / R21HD094051].

Notes on contributors

Marisa Meyer, B.A. is a postbaccalaureate research fellow at the National Institute of Mental Health (NIMH). She studied psychology at the University of Michigan, graduating head of her class with Highest Distinction. Marisa conceptualized and conducted an honors thesis, receiving the W.B.

Pillsbury Prize for outstanding research in the field of psychology. Marisa is currently refining her research and clinical skills through the NIMH Intramural Research Training Award.

Jennifer M. Zosh, Ph.D., is an Associate Professor of Human Development and Family Studies at Penn State University's Brandywine campus. As the Director of the Brandywine Child Development Lab, she studies early cognitive development. Most recently, her work has focused on the topic of playful learning in digital and real-world contexts. She received her Ph.D. in Psychological and Brain Sciences from The Johns Hopkins University.

Caroline McLaren is an undergraduate at the University of Michigan, studying Movement Science and Intraoperative Neuromonitoring. Since 2018, she has been a research assistant in the Radesky Lab through the University of Michigan Medical School.

Michael Robb is senior director of research at Common Sense. He has been involved in issues involving children and media for over 20 years, and has published research on the impact of digital media on young children's language development, early literacy outcomes, and problem-solving abilities in a variety of academic journals. Michael received his B.A. from Tufts University and M. A. and Ph.D. in psychology from UC Riverside.

Harlan McCaffery earned his MS in Biostatistics from Northwestern University in 2018, and has worked at the University of Michigan Department of Pediatrics since 2018. He has expertise in statistical programming, data visualization, and advanced statistical methods including mixed effects modeling, latent growth curve modeling, and survival analysis. He has collaborated on research related to infant growth and development, feeding and eating behaviors, and predictors of hospital length of stay.

Roberta Michnick Golinkoff, Ph.D., Cornell University, is known for her research on language development, the benefits of play, children's spatial learning, and the effects of media on children, with her research funded by Federal agencies and the LEGO Foundation. Playful Learning Landscapes, her latest project, marries architectural design and the science of learning to invite informal learning. Her last book, *Becoming Brilliant* reached the New York Times best-seller list.

Kathy Hirsh-Pasek, Professor of Psychology at Temple University and senior fellow at the Brookings Institution has written 16 books and 250+ publications. She served as President of the International Congress for Infant Studies and is on the Governing Board of the Society for Research in Child Development. Winning awards from every psychological and educational society for basic and translational research, she is passionate about putting learning science in the hands of parents and educators.

Jenny Radesky, M.D., Assistant Professor of Pediatrics at the University of Michigan Medical School, is a practicing developmental-behavioral pediatrician and media researcher. Her research examines design affordances of modern technology in the context of parent-child interaction and child social-emotional development. She authored the American Academy of Pediatrics policy statements on early childhood media use (2016) and digital advertising to children (2020).

ORCID

Marisa Meyer  <http://orcid.org/0000-0002-6837-7192>

Jennifer M. Zosh  <http://orcid.org/0000-0002-1698-4708>

Kathy Hirsh-Pasek  <http://orcid.org/0000-0003-2947-4544>

References

AAP Council on Communications and Media. (2016). Media and young minds. *Pediatrics*, 138(5), e20162591. Article doi: [10.1542/peds.2016-2591](https://doi.org/10.1542/peds.2016-2591).

- Barr, R. (2013). Memory constraints on infant learning from picture books, television, and touchscreens. *Child Development Perspectives*, 7(4), 205–210. doi:[10.1111/cdep.12041](https://doi.org/10.1111/cdep.12041)
- Barr, R., Shuck, L., Salerno, K., Atkinson, E., & Linebarger, D. L. (2010). Music interferes with learning from television during infancy. *Infant and Child Development*, 19(3), 313–331. doi:[10.1002/icd.666](https://doi.org/10.1002/icd.666)
- Binns, R., Lyngs, U., Van Kleek, M., Zhao, J., Libert, T., & Shadbolt, N. (Eds.) (2018). Third party tracking in the mobile ecosystem. *Proceedings of the 10th ACM Conference on Web Science*. Amsterdam, Netherlands: ACM Digital Library. [10.1145/3201064.3201089](https://doi.org/10.1145/3201064.3201089)
- Bond, B. J., & Calvert, S. L. (2014). A model and measure of US parents' perceptions of young children's parasocial relationships. *Journal of Children and Media*, 8(3), 286–304. doi:[10.1080/17482798.2014.890948](https://doi.org/10.1080/17482798.2014.890948)
- Callaghan, M. N., & Reich, S. M. (2018). Are educational preschool apps designed to teach? An analysis of the app market. *Learning, Media and Technology*, 43(3), 280–293. doi:[10.1080/17439884.2018.1498355](https://doi.org/10.1080/17439884.2018.1498355)
- Chiong, C., & Shuler, C. (2010). Learning: Is there an app for that? Investigations of young children's usage and learning with mobile devices and apps. The Joan Ganz Cooney Center at Sesame Workshop. <http://joanganzcooneycenter.org/publication/learning-is-there-an-app-for-that/>
- Christakis, D. A., Garrison, M. M., Herrenkohl, T., Haggerty, K., Rivara, F. P., Zhou, C., & Liekweg, K. (2013). Modifying media content for preschool children: A randomized controlled trial. *Pediatrics*, 131(3), 431–438. doi:[10.1542/peds.2012-1493](https://doi.org/10.1542/peds.2012-1493)
- Committee on Developments in the Science of Learning. (2000). *How people learn: Brain, mind, experience, and school: Expanded edition*. Washington, DC: The National Academie Press. [10.17226/9853](https://doi.org/10.17226/9853)
- Falloon, G. (2013). Young students using iPads: App design and content influences on their learning pathways. *Computers & Education*, 68, 505–521. doi:[10.1016/j.compedu.2013.06.006](https://doi.org/10.1016/j.compedu.2013.06.006)
- Fisher, A. V., Godwin, K. E., & Seltman, H. (2014). Visual environment, attention allocation, and learning in young children: When too much of a good thing may be bad. *Psychological Science*, 25(7), 1362–1370. doi:[10.1177/0956797614533801](https://doi.org/10.1177/0956797614533801)
- Hiniker, A., Lee, B., Kientz, J. A., & Radesky, J. S. (Eds.). (2018). Let's play! digital and analog play between preschools and parents. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Montreal, Canada: ACM Digital Library. <https://doi.org/10.1145/3173574.3174233>
- Hirsh-Pasek, K., Zosh, J. M., Golinkoff, R. M., Gray, J. H., Robb, M. B., & Kaufman, J. (2015). Putting education in "educational" apps: Lessons from the science of learning. *Psychological Science in the Public Interest*, 16(1), 3–34. doi:[10.1177/1529100615569721](https://doi.org/10.1177/1529100615569721)
- Huber, B., Tarasuik, J., Antoniou, M. N., Garrett, C., Bowe, S. J., Kaufman, J., & Team, S. B. (2016). Young children's transfer of learning from a touchscreen device. *Computers in Human Behavior*, 56, 56–64. doi:[10.1016/j.chb.2015.11.010](https://doi.org/10.1016/j.chb.2015.11.010)
- Kannass, K. N., & Colombo, J. (2007). The effects of continuous and intermittent distractors on cognitive performance and attention in preschoolers. *Journal of Cognition and Development*, 8(1), 63–77. doi:[10.1080/15248370709336993](https://doi.org/10.1080/15248370709336993)
- Linebarger, D., Barr, R., Lapierre, M., & Piotrowski, J. (2014). Associations between parenting, media use, cumulative risk, and children's executive functioning. *Journal of Developmental and Behavioral Pediatrics*, 35(6), 367–377. doi:[10.1097/DBP.000000000000069](https://doi.org/10.1097/DBP.000000000000069)
- Linebarger, D. N., Brey, E., Fenstermacher, S., & Barr, R. (2017). What makes preschool educational television educational? A content analysis of literacy, language-promoting, and prosocial preschool programming. In R. Barr & D. Linebarger (Eds.), *Media exposure during infancy and early childhood* (pp. 97–133). New York, NY USA: Springer. doi:[10.1007/978-3-319-45102-2_7](https://doi.org/10.1007/978-3-319-45102-2_7)
- Mayer, R. E. (2014). Research-based principles for designing multimedia instruction. In V. A. Benassi, C. E. Overson, & C. M. Hakala (Eds.), *Applying science of learning in education: Infusing psychological science into the curriculum* pp.(59–70). Washington, DC: Retrieved from the Society for the Teaching of Psychology website <http://teachpsych.org/ebooks/asle2014/index.php>
- Meyer, M., Adkins, V., Yuan, N., Weeks, H. M., Chang, Y. J., & Radesky, J. (2019). Advertising in young children's apps: A content analysis. *Journal of Developmental and Behavioral Pediatrics*, 40(1), 32–39. doi:[10.1097/DBP.0000000000000622](https://doi.org/10.1097/DBP.0000000000000622)

- Munzer, T. G., Miller, A. L., Weeks, H. M., Kaciroti, N., & Radesky, J. S. (2019). Differences in parent-toddler interactions with electronic versus print books. *Pediatrics*, 143(4), Article, e20182012. [10.1542/peds.2018-2012](https://doi.org/10.1542/peds.2018-2012).
- Outhwaite, L. A., Faulder, M., Gulliford, A., & Pitchford, N. J. (2019). Raising early achievement in math with interactive apps: A randomized control trial. *Journal of Educational Psychology*, 111(2), 284–298. <https://doi-org/10.1037/edu0000286>
- Radesky, J. S., Eisenberg, S., Kistin, C. J., Gross, J., Block, G., Zuckerman, B., & Silverstein, M. (2016). Overstimulated consumers or next-generation learners? Parent tensions about child mobile technology use. *Annals of Family Medicine*, 14(6), 503–508. doi:[10.1370/afm.1976](https://doi.org/10.1370/afm.1976)
- Radesky, J. S., Weeks, H. M., Ball, R., Schaller, A., Yeo, S., Durnez, J., ... Barr, R. (2020). Young children's use of smartphones and tablets. *Pediatrics*, 146(1), Article, e20193518. [10.1542/peds.2019-3518](https://doi.org/10.1542/peds.2019-3518).
- Rideout, V. (2017). The common sense census: Media use by kids age zero to eight. common sense media. <https://www.common SenseMedia.org/research/the-common-sense-census-media-use-by-kids-age-zero-to-eight-2017>
- Tare, M., Chiong, C., Ganea, P., & Deloache, J. (2010). Less is more: How manipulative features affect children's learning from picture books. *Journal of Applied Developmental Psychology*, 31(5), 395–400. doi:[10.1016/j.appdev.2010.06.005](https://doi.org/10.1016/j.appdev.2010.06.005)
- Vaala, S., Ly, A., & Levine, M. H. (2015). Getting a read on the app stores: A market scan and analysis of children's literacy apps. The Joan Ganz Cooney Center at Sesame Workshop. Retrieved from <http://joanganzcooneycenter.org/publication/getting-a-read-on-the-app-stores-a-market-scan-and-analysis-of-childrens-literacy-apps/>
- Vosniadou, S. (2001). *How children learn*. International Academy of Education. <http://www.ibe.unesco.org/>
- Yogman, M., Garner, A., Hutchinson, J., Hirsh-Pasek, K., & Golinkoff, R. M. (2018). The power of play: A pediatric role in enhancing development in young children. *Pediatrics*, 142(3), Article, e20182058. [10.1542/peds.2018-2058](https://doi.org/10.1542/peds.2018-2058).
- Zero to Three. (2018, October 25). *Choosing media content for young children using the E-AIMS model*. <https://www.zerotothree.org/resources/2533-choosing-media-content-for-young-children-using-the-e-aims-model#downloads>

Appendices

Appendix A. pillar 1 – Active Learning.

Score	Description
0	<p>Pillar objectives not met. App includes:</p> <ul style="list-style-type: none"> • Activities that can be completed with little mental effort; and/or • Activities that elicit simple “stimulus-response reactions to on-screen actions;”¹ and/or • Passively following instructions with extremely constrained design that involves tapping the screen to make pages advance.
1	<p>Activities have some learning objective, but are structurally closed-loop or repetitive.</p> <ul style="list-style-type: none"> • There are many hints, demonstrations, or prompts to tell the child what to do; and/or • App drills skills, and feels like an online workbook; and/or • App forces activities to automatically advance from one activity to the next.
2	<p>Activities have a learning objective, with some degree of flexibility or child-led problem-solving to complete them.</p> <ul style="list-style-type: none"> • There is a mild degree of mental challenge to the app activities, such as having the player actively follow the programmed instructions or narrations to complete the learning objective. • Choices for completing the task may be constrained.
3	<p>Pillar objectives met. App is a “minds-on” game, and includes activities “that require thinking and intellectual manipulation.”¹</p> <ul style="list-style-type: none"> • Activities allow the player to generate responses and incrementally provide more challenge. • There may be several steps for the child to follow; however, the steps aren’t spoon-fed to them, and hints are subtle or appropriate.

Appendix A. pillar 2 – Engagement in the Learning Process instructions: treating each of the three category requirements as “ANDs,” start assessing the app as if it earns a 3, and then move down the categories until all three “ANDs” are met.

Score	Distracting Advertising	Distracting Feedback	Gamification
0	<ul style="list-style-type: none"> • Embedding ad videos in gameplay; and/or • If pop-up ad interrupts mid-gameplay (e.g., in the middle of making a dessert, but not after completing one puzzle – must interrupt gameplay) more than one time; and/or • 4+ disruptive advertising present. 	<ul style="list-style-type: none"> • <u>Object feedback</u> provided after each action (e.g., <u>earning a coin</u> and seeing it put in the piggy bank and hearing a “cha-ching” noise after each step in an activity – like after spraying water on a shirt to be ironed). 	<ul style="list-style-type: none"> • No active learning. Design features are purely for entertainment purposes.
1	<ul style="list-style-type: none"> • 1-3 kinds of disruptive advertising present (includes pop-up ads, banner ads, hidden ads, rating prompts, social media prompts, and in-app purchases). 	<ul style="list-style-type: none"> • Feedback at the end of the game and/or after each action. • 6+ visual or auditory feedback mechanisms, and 1+ object reward. 	<ul style="list-style-type: none"> • Extraneous visual and sound effects, which may distract the user from the game. • (No code 2).
2	<ul style="list-style-type: none"> • Only in-app purchases allowed to be present. • No disruptive advertising should be present. • If character promotes in-app purchases, give a 0 code for emotional manipulation. 	<ul style="list-style-type: none"> • Feedback at the end of the game only. • 4–5 types of feedback provided, of which only 1 should be an object reward. • If more than one object reward (potential to earn coins and a sticker), give a 1 code. 	
3	<ul style="list-style-type: none"> • No advertising present. 	<ul style="list-style-type: none"> • Feedback at the end of the game only. • <u>Very minimal</u> – 3 or less types of feedback, of which only 1 should be an object reward. 	<ul style="list-style-type: none"> • Visual and sound effects included only to enhance learning and engagement in the game.

Appendix A. pillar 3 – Meaningful Learning.

Score	Description
0	Pillar objectives not met. App's activities have no relevance to child's life experiences, and no learning objectives.
1	App activities involve rote learning. The learning objectives are relevant to the child's life, but the design of the activities and the presentation of the learning objectives do not promote meaningful learning.
2	App activities go beyond rote learning and repetition, and provide opportunities for meaningful learning in a digital setting; however, the app doesn't explicitly prompt the child to engage in similar activities outside of their device, or these prompts are subtle and wouldn't be easy to follow if the user can't read. <ul style="list-style-type: none"> • App activities may attempt to link prior knowledge to new content. • App activities may model new ways of thinking or acting. • For apps that promote development of rote skills (such as <i>Starfall ABC</i> or <i>Endless Alphabet</i>), they must provide multiple examples of a concept. Activities and explanations must go beyond a simplistic presentation of "this is X," and instead discuss or demonstrate how the concept is contextually meaningful. • Ask yourself: "does the app ask the child to go beyond rote learning?"
3	Pillar objectives met. <ul style="list-style-type: none"> • App's activities based on meaningful learning experiences that would be easy for the child to then do in real life. • High-quality connections between gameplay and the child's life. • Gameplay narrative and principal learning objective strongly intertwined.

Appendix A. Pillar 4 – Social Interaction.

Score	Description
0	Pillar objectives not met. <ul style="list-style-type: none"> • App structure does not provide opportunities for in-person interactions or mediated interactions, and • App structure does not provide opportunities for development of parasocial relationships with characters in the app.
1	<ul style="list-style-type: none"> • App may have features to involve parents or another player, but also has lots of gamification that would grab child's attention and make it hard to get responsive interactions with the child; and/or • Potential for development of weak parasocial relationship with app characters.
2	<ul style="list-style-type: none"> • There is a clear space for a parent or sibling to play along (without a prompt to do activities with that person in the real world), such as through prompts for the user to engage in turn-taking or multi-touch input on-screen that invites others to play; and/or • App provides opportunity for development of high-quality parasocial relationships with app characters.
3	Pillar objectives met. <ul style="list-style-type: none"> • App structure provides opportunities for social interaction, including "face-to-face interactions around the screen,"¹ such as the parent playing together with the child; and • Ideas are presented for how to transfer knowledge to the real world with a parent or sibling.

Appendix B. Sampled apps and associated pillar scores.

App Name	Pillar 1 (Active Learning) Score	Pillar 2 (Engagement in Learning Process) Score	Pillar 3 (Meaningful Learning) Score	Pillar 4 (Social Interaction) Score	App Total Score
Duolingo: Learn Languages Free	1	0	2	0	3
Children's Doctor Dentist	1	1	1	1	4
ABC Kids – Tracing & Phonics	1	1	1	1	4
ABC Preschool Free	2	1	1	0	4
LEGO DUPLO Town	3	3	2	1	9
Baby Puzzles	1	1	1	0	3
Baby Panda's Potty Training – Toilet Time	1	1	2	1	5
Little Panda Travel Safety	1	0	3	1	5
Bible App for Kids: Read the Nativity Story	2	3	1	1	7
Colors for Kids, Toddlers, Babies – Learning Game	1	1	1	0	3
Sweet Baby Girl Daycare	1	0	1	1	3
Animals Puzzle for Kids	1	1	1	0	3
Sweet Baby Girl – Dream House and Play Time	1	0	1	1	3
Rosetta Stone: Learn to Speak & Read New Languages	1	1	2	0	4
Draw.ly – Color by Number Pixel Art Coloring	2	0	2	0	4
Dentist games for kids	1	1	1	1	4
Baby Panda Home Safety	1	0	3	1	5
McPlay	2	1	1	2	6
Baby Phone – Games for Babies, Parents and Family	0	0	0	0	0
Kids Educational Game 5	1	1	1	2	5
Little Panda Earthquake Safety	1	0	3	2	6
Super ABC! Learning games for kids! Preschool apps	1	1	1	0	3
Kids Learn Professions	1	1	1	0	3
Educational Games 4 Kids	1	1	1	2	5
Unir Puntos – Animales	1	1	1	0	3
Farming Simulator 18	3	0	2	2	7
Toca Lab: Elements	2	3	2	1	8
Teach Your Monster to Read – Phonics and Reading	1	1	1	1	4
Stellarium Mobile Sky Map	2	3	2	0	7
Drawing for Kids and Toddlers	2	0	2	0	4
Learn to Read with Tommy Turtle	1	1	1	1	4
Star Walk – Night Sky Guide: Planets and Stars Map	2	3	2	0	7
XtraMath	1	1	1	2	5
Stack the States	1	1	1	0	3
Vocabulary.com	1	1	2	0	4
Sight Words Games & Flash card	1	1	1	1	4
Kids Learn About Animals	1	1	1	0	3
Solar Walk 2 – Spacecraft 3D and Space Exploration	2	2	2	0	6
Learn Letter Sounds with Carnival Kids	1	1	1	0	3
Learn Letter Names and Sounds with ABC Trains	1	1	1	0	3
Kids ABC Tracing and Alphabet Writing	1	1	1	0	3
Daniel Tiger's Stop & Go Potty	2	3	3	2	10
Second Grade Learning Games	1	1	1	0	3

(Continued)

Appendix B. (Continued).

App Name	Pillar 1 (Active Learning) Score	Pillar 2 (Engagement in Learning Process) Score	Pillar 3 (Meaningful Learning) Score	Pillar 4 (Social Interaction) Score	App Total Score
Zoombinis	3	3	2	1	9
Stack the Countries	1	1	1	0	3
PAW Patrol Rescue Run	0	1	1	2	4
Peekaboo Barn	0	3	1	0	4
Sago Mini Pet Cafe	1	3	1	1	6
BRIO World – Railway	3	2	2	1	8
Monkey Preschool Fix-It	1	1	1	1	4
Math Learner: Easy Mathematics	1	1	1	1	4
HOMER Reading: Learn to Read	1	1	2	1	5
Kiddopia – ABC Toddler Games	1	1	1	1	4
Hidden Pictures Puzzles	1	1	2	1	5
Epic!	1	1	1	2	5
NOGGIN Preschool	0	3	3	2	8
SkyView Lite	3	1	3	0	7
PBS KIDS Video	0	1	3	2	6
Quizizz: Play to Learn	1	1	1	1	4
busuu – Language Learning	1	1	2	0	4
My Very Hungry Caterpillar AR	2	3	2	1	8
Nick Jr.	0	1	0	2	3
PBS KIDS Games	1	1	1	2	5
Speech Blubs: Language Therapy	1	1	2	1	5
Main Street Pets Big Vacation	2	0	2	1	5
Dr. Panda Train	1	1	1	1	4
Toca Life: World	3	2	2	1	8
TinyTap – Educational Games	1	1	1	0	3
Grades K-5 Math Learning Games	1	0	1	1	3
Drawing for kids: doodle games	2	1	1	0	4
Learn Languages with Memrise	1	1	1	0	3
Endless Learning Academy	1	2	2	1	6
Fiete Sports Games for Kids	1	1	1	1	4
Busy Shapes & Colors	1	2	1	0	4
Baby Piano for Kids & Toddlers	0	1	0	1	2
Toca Hair Salon 3	2	3	2	1	8
Toca Kitchen 2	2	3	2	1	8
Toca Life: Neighborhood	3	3	2	1	9
Toca Life: School	3	3	2	1	9
Toca Life: Hospital	3	3	2	1	9
Elmo Loves ABCs	1	1	2	2	6
Endless Alphabet	1	1	2	1	5
Monkey Preschool Lunchbox	1	1	1	1	4
My Town: Airport	2	1	2	1	6
Preschool All In One Basic Skills	1	1	1	0	3
Counting & Numbers. Learnin ...	1	1	1	1	4
Teenage Mutant Ninja Turtles: Half-Shell Heroes	0	0	0	1	1
Phonics Genius	1	3	1	0	5
Metamorphabet	1	3	2	0	6
Melody Jams	2	3	2	1	8
Tongo Music – for kids	0	3	1	1	5
My Food – Nutrition for Kids	2	3	3	2	10
Montessori Numberland	1	1	1	1	4
Curious George Train Adventures	1	1	1	1	4
Grandma's Preschool	1	1	1	1	4
Eric Carle's Brown Bear Animal Parade	1	3	1	1	6

(Continued)

Appendix B. (Continued).

App Name	Pillar 1 (Active Learning) Score	Pillar 2 (Engagement in Learning Process) Score	Pillar 3 (Meaningful Learning) Score	Pillar 4 (Social Interaction) Score	App Total Score
Sago Mini Puppy Preschool	1	1	1	1	4
Super Why! Phonics Fair	1	1	1	2	5
Montessorium: Intro to Letters	1	1	1	0	3
The Orchard by HABA	1	1	1	1	4
LEGO DUPLO Train	1	3	1	1	6
LEGO Juniors Create & Cruise	1	1	1	1	4
Colors & Shapes – Kids Learn Color and Shape	1	1	1	1	4
Balloon Pop Kids Learning Game Free for babies	0	0	0	0	0
ABC Alphabet games for toddlers! Learning letters!	1	0	2	0	3
Spanish Lucas' Whiteboard	2	1	1	0	4
Monster Trucks Game for Kids 2	1	0	1	1	3
123 Numbers – Count & Tracing	1	1	1	1	4
Sago Mini Holiday Trucks and Diggers	2	3	2	1	8
Sago Mini Friends	2	3	2	1	8
Baby Panda's Fashion Dress Up Game	2	1	1	1	5
ABC Animals	1	1	1	1	4
Baby games: puzzles for kids	1	1	1	0	3
Toddler Learning Game – EduKitty	1	1	1	1	4
KidsDoodle	2	1	2	0	5
123 Toddler games for 2+ years	1	1	1	0	3
Learning Games 4 Kids Toddlers	1	1	1	1	4
Match games for kids toddlers	1	1	1	0	3
PaintSparkle	2	1	2	0	5
Preschool All-In-One	1	1	1	1	4
Sorter Deluxe	1	1	1	0	3
Toddler Games: puzzles, shapes	1	1	1	0	3
Tozzle Lite – Toddler's favorite	1	1	1	0	3
The Wheels on the Bus Musical	1	1	1	0	3