Can apps support creativity in middle childhood?

Piotrowski, J.T.; Meester, L.

DOI
10.1016/j.chb.2018.03.030
Publication date
2018
Document Version
Final published version
Published in
Computers in Human Behavior

Computers in Human Behavior 85 (2018) 23–33

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Full length article

Can apps support creativity in middle childhood?

Jessica Taylor Piotrowski [a, *], Laurian Meester [b]

[a] Amsterdam School of Communication Research ASCoR, University of Amsterdam, Amsterdam, The Netherlands
[b] Graduate School of Communication, University of Amsterdam, Amsterdam, The Netherlands

ARTICLE INFO

Article history:
Received 22 August 2017
Received in revised form
15 March 2018
Accepted 19 March 2018
Available online 20 March 2018

Keywords:
Apps
Children
Creativity
Engagement
Middle childhood

ABSTRACT

This experimental study investigated whether and how creative apps may support creativity in middle childhood (n = 94 children, 8–10 years old). Guided by the moderate discrepancy hypothesis, flow theory, and the differential susceptibility to media effects model, developmentally-appropriate creative apps were predicted to increase engagement and subsequent creativity to a greater extent than developmentally-inappropriate creative apps. Furthermore, gender and fantastical thinking were predicted to moderate effects. Results provided partial support for study hypotheses. Children were more engaged when playing developmentally-appropriate apps, however, this engagement did not translate into creativity gains. Given that the data failed to support study hypotheses, post hoc analyses were conducted to explore the findings in greater detail. These additional analyses indicated that developmentally-appropriate apps not only lead to greater engagement, but were subsequently more appealing as well. These post hoc findings are discussed within the context of the study design, particularly noting that greater duration of play may be necessary to move the needle on creative skills. Empirical and practical implications are discussed.

With more than one million apps produced in today's App Stores, apps have become of great interest in today's digital society. Nearly 80% of American families already have at least one tablet in their household in 2014 (Rideout, 2014). Although this number is somewhat lower in Western European industrialized countries, this still reflects a substantial number of children that come into contact with apps. Interestingly, having children is one of the main reasons that parents mention for purchasing tablets (Maytom, 2014), and indeed, the estimates suggest that children are coming into contact with apps at younger and younger ages. In fact, recent data with American children suggests that young children's access to mobile technology may be approaching saturation (Kabali et al., 2015) while similar patterns have been increasingly noted throughout the industrialized world (Holloway, Green, & Livingstone, 2013). These so-called digital natives have made the digital language their own since they were born (Prensky, 2001).

When it comes to apps for children, sales data across the United States, Germany, Spain and the Netherlands show that educational apps are among the most downloaded (paid) apps. Defined as apps which are designed to support a range of positive outcomes related to literacy, mathematics, social skills, and science (Chiong & Shuler, 2010; Hirsh-Pasek et al., 2015), thus far, the burgeoning literature indicates that educational apps with clear curriculum goals do seem to work (e.g., Chiong & Shuler, 2010; Krcmar, 2010) – particularly among younger children. Yet, the popularity of such didactically-focused apps dramatically decreases as children grow older. By middle childhood, these apps are firmly replaced in popularity with apps which encourage creativity (Dickson, 2014).

Middle childhood, a period of time representing children 8–11 years of age, is characterized by rapid cognitive and emotional development (Valkenburg & Cantor, 2000) – including a more logical thought process and greater interest in identifying and solving problems (Valkenburg & Piotrowski, 2017), Given this growing interest in problem-solving, it is perhaps not that surprising that children this age are gravitating towards apps which aim to support creativity. Creativity is defined as the ability to apply new ideas to an open-ended task (Amabile, 2012) which are meaningful and surprising (McPake, Plowman, & Stephen, 2013). Jackson et al. (2012) further explain that creativity is a mental process in which new ideas or new connections between existing ideas are generated. This mental process reflects an interplay between defining, making, and perceiving in which problem-solving plays a key role (Burnard & Younker, 2004). Considering the fact that children in middle childhood are spending a good deal of their time with tablet-based media (Rideout, 2014), particularly creative-

* Corresponding author. Amsterdam School of Communication Research, University of Amsterdam, PO Box 15791, 1001 NG Amsterdam, The Netherlands.
E-mail address: j.piotrowski@uva.nl (J.T. Piotrowski).

https://doi.org/10.1016/j.chb.2018.03.030
0747-5632/

based apps (iOS Top Apps Charts, 2015), and that they are developing creative skills such as problem-solving, it is reasonable to ask whether and how these apps might support children's development of creativity. Yet, to date, there exist no studies investigating apps and children's creativity. To address this gap, we use an experimental design with children aged 8–10 years old to evaluate whether creative apps support creativity in middle childhood.

1. Media & creativity

Although children in middle childhood exhibit interest in creative-based media content, the empirical literature has largely ignored the potential opportunities of such media for supporting creative skills. This is somewhat surprising given that most studies on creativity show that adults who are creative tend to come from families that have provided them with a favorable background for the development of intellectual abilities (Mumford & Gustafson, 1988). And, as argued by Valkenburg and Piotrowski (2017), if environmental forces during childhood can influence the development of creativity, it is reasonable to expect that media may be one such socializing factor with a great potential to influence this development. Yet, while there is a sizeable body of work which has investigated the potential perils of media for creativity, the opportunities are far less understood.

Early work on media and creativity (Runco & Pezdek, 1984) suggested that neither television nor radio supports children's creativity. Somewhat later, Valkenburg and Beentjes (1997) offered a more nuanced perspective. In their work, they showed that – among children 8–10 years of age - exposure to radio supported greater creativity (as indexed by the number of novel ideas) when compared to exposure to television, whereas this effect was not present for younger children. And more recently, in their 2012 review of the literature on media and creativity, Valkenburg & Calvert found evidence to suggest that children's educational television programs such as Mister Rogers Neighborhood, Barney, and Dora the Explorer (i.e., content which included creativity as a curriculum aim) can bolster young children's store of ideas and subsequently enrich their imaginative play and creative tasks. This "stimulation hypothesis" argues that young viewers encode events from the media (e.g., characters, settings, scenes) and then transform and incorporate this information into their own play patterns (Valkenburg & Piotrowski, 2017).

Thus, it seems that – over time – the scholarship on media and creativity in childhood has slowly taken a more positive view. And yet, still the field remains thin. Indeed, despite a quickly growing digital media space, there are nearly no studies about the opportunities of apps – or any form of digital media – on children's creativity. One exception is work by Jackson et al. (2012). These scholars found that, for 12-year-olds, greater video game playing (of all types) was associated with greater creativity as measured by the figural subtest of the Torrance test, even after adjusting for demographic differences. Yet direction of effect was unclear. And more generally, meta-analytic work (Powers, Brooks, Aldrich, Palladino, & Alfieri, 2013) suggest that, regardless of genre, gaming may support fluid intelligence (i.e., one's ability to think logically and solve problems in novel situations, independent of acquired knowledge; Flynn, 1999) – often argued to be a predictor of creativity (Nusbaum & Silvia, 2011). The argumentation, while underdeveloped, is that gaming encourages cognitive practice which – over time – strengthens ones' fluid intelligence which in turn can lead to more divergent thinking patterns.

Compared to scholarship in other fields of media effects such as educational television effects or serious gaming effects, the connection between digital media and creativity has been largely underexplored. This notable absence is particularly surprising when one considers the numerous opportunities for young users to create, experiment, and construct in the interactive space. All told, we see evidence from television research that suggests a positive link, but simultaneously points to the importance of *developmentally-appropriate* content designed with the intent to support creativity. The evidence from digital media research, while markedly smaller, also gives some reason for hope. Merging these two bodies of work in the context of today's ever-popular app world, it seems fair to suggest that creative apps can support creativity in childhood. However, the boundary conditions of such effects – particular in terms of both the content and the user - are even less understood.

2. Theoretical boundary conditions

2.1. Content boundaries

Since the minimal existing work on media and creativity provides little insight as to when creative content (namely, apps) are expected to benefit children's creative skills, it is reasonable to turn to the literature on children's media preferences to help identify the types of creative content which may be most effective in supporting this goal. The moderate discrepancy hypothesis (MDH; Valkenburg & Cantor, 2000) states that children prefer media content which is only moderately discrepant from their capabilities. This hypothesis has been confirmed for television, for example, with researchers demonstrating that children prefer watching content which they can easily understand, but avoid content for which they do not yet have the capabilities to fully comprehend the content (Valkenburg & Vroone, 2004). It offers one of the most plausible arguments for why media preferences are developmentally-dependent.

Although the MDH has not yet been tested for apps, there is some evidence to indicate that the MDH can be used within this content. In her research with apps, Aziz (2013) found that 5-year-olds had different preferences in app content and design than 8-year-olds such that content which did not match developmental skills resulted in boredom and distraction. Similarly, consistent with the MDH, Broekman and colleagues (Broekman, Piotrowski, Beentjes, & Valkenburg, 2016) found that – when it comes to selecting apps –parents of younger children express a greater need for paced, recognizable and controllable content, while parents of older children express a greater need for more complex content.

In terms of effective creative apps in middle childhood, the MDH would thus expect that apps which differ only moderately from what the user's developmental capabilities will be more enjoyable and ultimately more effective. This proposition is line with propositions of flow theory (Csikszentmihalyi, 1997) which suggests that digital content will be most enjoyable and engaging when users achieve a state of flow (Sherry, 2004). This flow state occurs when cognitive skills of the person and the task of the game are in balance. Considering the fact that children's learning from media is typically predicated upon the importance of them enjoying the media content (Vorderer, Klimmt, & Ritterfeld, 2004), it is reasonable to argue that content which is moderately discrepant from what they know will help induce flow which ultimately will support game enjoyment and learning. As such, we expect that moderately discrepant creative apps will support children's creativity to a greater extent than highly discrepant creative apps (H1).

Hypothesis 1. Moderately discrepant apps will support children's creativity to a greater extent than highly discrepant apps.

2.2. User boundaries

Both MDH and flow theory make the important point that the

successful balance between content and user skills will predict learning. Flow theory, in particular, highlights the role of the media response *process* in which the user is engaged with the media content, which subsequently induces a feeling of flow resulting in enjoyment. Flow theory is not, however, the only theory to suggest that how we respond to media content predicts its effects. There are a host of media effects theories which similarly suggest that media processing is critical towards understanding media effects (e.g., social learning (cognitive) theory; Bandura, 1977; elaboration likelihood model; Cacioppo, Petty, & Stoltenberg, 1985). Most recently, the Differential Susceptibility to Media effects Model (DSMM) (Valkenburg & Peter, 2013) has joined this list of media effect theories. Built upon a range of media effects theories, the model posits that the user's cognitive, excitative, and emotional responses to media content predict whether and how media will affect its users. In other words, the DSMM expects that media responses mediate the relationship between media content and media effects. When it comes to digital media content, flow theory suggests that these responses can be operationalized as engagement (Hamlen, 2011).

Although varied definitions of engagement exist, the majority of researchers argue that engagement consists of cognitive, behavioral, and emotional reactions to media content (Annetta, Minogue, Holmes, & Cheng, 2009; Dickey, 2005). Arguably akin to the DSMM's response states, engagement is said to be a key indicator of game involvement (Brockmyer et al., 2009) and predictive of flow (Hamlen, 2011). Research with digital media (particularly video games) has shown that engagement is a key mediator in the relationship between media exposure and effects. Users who experience increased engagement during game play experience enhanced learning effects (Howard-Jones & Demetriou, 2009; Huizenga, Admiraal, Akkerman, & Dam, 2009). Given that the DSMM predicts that heightened behavioral, cognitive, and excitative responses to media content are expected to strengthen media effects, and that research has shown that engagement induces a state of flow, we expect that moderately discrepant apps will lead to increased engagement and subsequent increased creativity (H2).

Hypothesis 2. Engagement will mediate the relationship between apps and creativity such that moderately discrepant apps will lead to increased engagement and subsequently increased creativity.

While both the DSMM and flow theory note that *how* users' process media will play a key role in influencing media effects, the DSMM further argues that this process may be differentially affected by individual differences. Specifically, the model argues that there are unique dispositional (e.g. gender, personality, values), developmental (e.g. cognitive, emotional development), and societal factors (e.g. friends, culture) which may impact a user's response to media content (Valkenburg & Peter, 2013). These variables can predict media use or moderate the influence of media exposure - leading to stronger or weaker effects. Although there are numerous variables which may affect how children process creative, the existing literature suggests that two variables are of particular interest: gender and fantastical thinking.

2.3. Gender

In terms of gender, several studies suggest that boys and girls engage with mediated content in differing ways. Boys tend to be more engaged with content that offers challenge and competition, whereas girls are less fond of competition (Inal & Cagiltay, 2007; Valkenburg & Piotrowski, 2017). Interestingly, the majority of creative apps tend to limit or avoid competition altogether, favoring instead an open approach in which they cannot fail (Cohen, Hadley, & Frank, 2011). It is perhaps not surprising then that girls tend to prefer creative content more than boys and actually search for this type of media content (Valkenburg & Piotrowski, 2017; Weber & Mitchell, 2008, pp. 25–47). Given their preference for creative content as well as their preference for non-competitive gaming (i.e., a key characteristic of many creative apps), it is reasonable to expect that girls may engage with creative apps to a greater extent than boys, and as a result, may benefit more from this content.

Hypothesis 3. Gender will moderate the relationship between apps, engagement, and creativity such that girls will be more engaged with moderately discrepant apps than boys and will subsequently experience greater creativity.

2.4. Fantastical thinking

Just as gender is expected to moderate how children process creative apps, the existing literature suggests that the extent to which children are prone to fantastical thinking may also influence this processing. Fantastical thinking (also referred to as fantasy) is a conscious state that occasionally takes place with one's attention shifting away from the tasks that one was initially doing or thinking about because images are evoked in one's head of certain situations that are not really happening (Singer 1966 in Valkenburg & Peter, 2006). Fantasy-prone children have a vivid imagination (Woolley, 1997), are more likely to be more creative (Mullineaux & Dilalla, 2009), and score better on divergent thinking tasks (Russ, Robins, & Christiano, 1999) which is an important skill linked to creativity (Runco & Okuda, 1988). These children typically enjoy activities which have limited boundaries (Dansky, 1980) and like to use their ability for detailed "make-believe" for entertainment. Since the majority of creative apps have minimal rules and instead encourage open-thinking and engagement (Cohen et al., 2011), it is possible that these apps are particularly engaging for children who are prone to fantasy, and as a result, may benefit more from this content.

Hypothesis 4. Fantastical thinking will moderate the relationship between apps, engagement, and creativity such that children with greater fantastical thinking will be more engaged with moderately discrepant apps and subsequently experience greater creativity.

3. The current study

It is clear that apps are becoming an important part of the everyday lives of today's children. While educational apps, in general, tend to decrease in popularity during middle childhood, we see that the popularity of creative apps tends to peak during this period. Researchers and educators now consider creativity an important 21st century learning skill, and yet surprisingly, this potential opportunity for creative apps to support this key skill has not yet been investigated. Understanding whether creative apps can support creativity is an important direction with both theoretical and applied consequences. To that end, using an experimental design that is guided by propositions of the MDH, flow theory, and the DSMM, this study investigates whether creative apps can support children's creative skills – with particular attention to the role of engagement, gender, and fantastical thinking in this process. Fig. 1 depicts the conceptual model guiding this study.

4. Method

4.1. Research design

To address study hypotheses, a between-subjects experiment

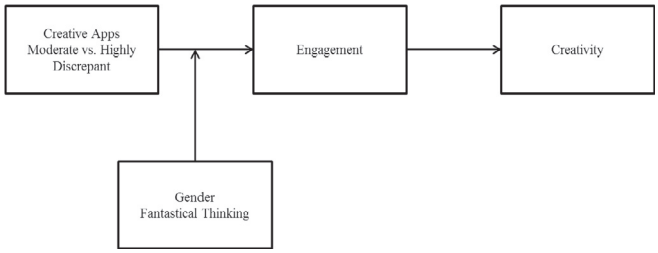

Fig. 1. Conceptual model.

was conducted. Children were randomly assigned to one of two conditions: highly discrepant apps ($n = 49$) and moderately discrepant apps ($n = 45$).

4.2. Participants

After receiving approval from the Institutional Review Board at the sponsoring institution, children were recruited at a children's media film festival in the Netherlands. Only children between the ages of 8–10 years of age with parental consent were eligible to participate. During the festival, one of two trained research assistants approached parents about the potential of their children participating in the study. In total, 94 children participated in the study (girls = 59.6%; $M_{age} = 9.52$ years, $SD = 0.92$). Participants were randomly assigned to one of the two conditions. Groups were not significantly different from each other in terms of age ($t(92) = -0.45$, $p = .654$, $X^2 = 34.43$, $p = .636$), gender ($\chi^2 = .01$, $p = .936$), or fantastical thinking ($t(92) = 1.00$, $p = .317$).

Of the 94 children, 50 parents completed a parent questionnaire designed to provide more detailed information about study participants. Parent questionnaire data indicated that the majority of parents worked at least part-time outside of the home (mothers: 76%; fathers: 86%), and that participating children were experienced with touchscreen technology. Parents reported that children, on average, spend approximately 30 min per day on digital devices such as tablets and smartphones, of which about a third of that time is spent playing with educational apps. This is similar to estimates found within the Netherlands and in other industrialized countries (Rideout, 2014).

4.3. Materials

To help ensure that effects were not related to a specific app, two apps that were designed to support children's creativity were selected for each condition. Selection occurred in several steps. First, the most popular app recommendation tools for parents were identified both within the Netherlands (where the study was conducted) and abroad, namely, the Cinekid AppLab app[1] (Dutch) and Common Sense Media (American).[2] Working with these tools, a pool of potential apps was developed for both the moderately discrepant and highly discrepant conditions.

First, working in the Cinekid AppLab, apps that were identified as supporting creativity (i.e., in "Arts and Crafts" category) and were coded as appropriate for children 8–10 years old (i.e., moderately discrepant) or appropriate for children 3–4 years old (i.e., highly discrepant) were listed in their respective pool. Then, working with Common Sense Media's 'best creative app' list, we selected those apps that targeted children 8–12 years old (i.e., moderately discrepant) and those apps that targeted children 2–4 years old (i.e., highly discrepant) and assigned them to their respective pool. In total, our selection yielded 47 potential apps for inclusion (i.e., 25 = moderate; 22 = high).

Once these potential apps were identified, the second step was to confirm that these apps were of high-quality and would meet the needs of our study. To assess quality, we conducted a literature search to identify what is most agreed upon as best features for games, digital media, and creativity. In total, we identified eleven criteria for which to code the apps. These criteria were: presence of clear goal, includes challenging task(s), provides corrective feedback, provides good and accessible instruction, is novel, includes a variety of choices, is simple to play, uses a blend of education and fun, has high production value, offers a task to accomplish, and provides a sense of control over actions (Dickey, 2005; Falloon, 2013; Garzotto, 2007; O'Hare, 2014; Pavlas, 2010). In addition, there were also four criteria specific to the study: app is not gender specific, app can be played within 10 min, app is understandable for Dutch speaking children, and the app contains a single-player mode.

All 47 apps were coded for the quality and study criteria by two independent researchers. Reliability statistics indicated acceptable intercoder reliability (Krippendorff's $\alpha = .75$). Apps that scored the highest on quality criteria as well as met the study criteria were eligible for inclusion. These eligible apps were then provided to four judges [scholars of children and media] where, based on the blinded app description, the judges were asked to identify whether they felt the app targeted younger children or children in middle childhood. There was no discrepancy in these responses – all mapping onto our expectation based on the app recommendation tools - adding confidence in our use of these tools as a starting point. At the final stage, we randomly selected 2 moderately discrepant apps (*Toca Hair Salon Me* and *Easy Studio*) and two highly discrepant apps (*Nick Jr. Draw & Play* and *Sago Mini Monster*) for study inclusion. A description of these apps from Common Sense Media (https://www.commonsensemedia.org/) is found in Table 1.

4.4. Measures

4.4.1. Fantastical thinking

Proneness to fantastical thinking was measured with an adapted scale from Rosenfeld, Huesmann, Eron, and Torney-Purta (1982). The original scale (titled the Children's Fantasy Inventory) consisted of 45 items. Published validity statistics indicated that the Fanciful-Intensity subscale of this assessment demonstrated the strongest validity. This subscale consisted of 23 items measuring vividness, scariness, fanciful thoughts, and intellectual concepts of fantasy. In order to prevent participant fatigue, a subset of these items was used for this study. Working with the published factor loadings in Rosenfeld et al. (1982), items that loaded higher than 0.45 were included in our scale ($n = 11$ items). An example item is "Have you ever have a make-believe friend who you talked to and who went to places with you?" All questions were answered using a 3-point Likert scale ("never", "sometimes", "often"). Items were averaged with higher scores indicating greater likelihood of fantastical thinking ($M = 1.64$, $SD = 0.35$, $\alpha = .74$).

4.4.2. Engagement

Engagement was measured with an adapted version of the Game Engagement Questionnaire (Brockmyer et al., 2009). Although the original scale consists of 19 items, only items that

[1] Developed primarily for use with Dutch parents, the Cinekid AppLab is a parental tool that provides reviews and recommendations for children's apps based on content and age categorization. Only apps that are deemed high-quality by a panel of AppLab reviewers are included in the AppLab.

[2] Similar to the Cinekid AppLab, Common Sense Media is a well-known American resource that provides parents with information and recommendations about a range of media including books, television shows, movies, and apps.

Table 1
App descriptions from common sense Media's "What's it about?"

Sago Mini Monsters	Sago Mini Monsters, From a green pit of bubbly slime, kids drag a monster shape. They choose from five colors to decorate their monster and tap the check mark to bring it to life. Horns pop into place; eyes and teeth appear. If they don't like what they see, though, kids can pull off the parts and new ones will grow into place. Then they feed and care for their monster, brushing its teeth and adding accessories. Kids can shake the device to knock everything off and start over at any time. When they're happy with their creation, they can take its picture.
Nick Jr. Draw & Play	Nick Jr Draw & Play begins with Dora the Explorer walking kids through the app's features and functionalities and instructing them how to tap and swipe. Kids select what type of material they'd like play on, then they are free to draw and play with their choice of instrument. Some of the tools offer special effects, such as magic wands for sparkling silver stars, swirling green leaves, flying black bats, etc. Kids can save their creations to an in-app gallery, where they can return to work on them later, or to the device's photo gallery.
Easy Studio Stop-Motion Studio	Easy Studio Stop-Motion Studio offers both an Easy Mode and an Expert Mode. In Easy Mode, a tutorial demonstrates how to move shapes, take photos, and move shapes again to create an animation. Five more lessons help kids learn basic techniques by creating an animated accordion, a train, an airplane, and more. Expert Mode provides all the same shapes and tools with no creative direction; kids simply press Play or tap into the Library to watch saved animations. In the Library, kids can edit animations or add an audio track in the form of their own recorded voice or pre-recorded silly noises and musical notes.
Toca Hair Salon Me	Tap "play" on Toca Hair Salon Me's main page and select a photo from your device's camera roll to style. Choose a head shape and pinch or expand your fingers to size the image to the head. Tap the green arrow to move onto the mouth- and eyes-sizing and centering page. When complete, start styling! Cut, curl, comb, color, and more by tapping and moving your finger around. Swipe the bottom of the screen for more options.

were developmentally appropriate for children aged 8–10 were selected ($n = 6$ items). Further, the wording of the selected questions was adjusted to reflect state engagement as opposed to trait engagement. The six items were (translated from Dutch): "during the games, I forgot the time", "while playing the games, the games seemed real", "I played without thinking how to play" [deleted due to low reliability], "playing the games gave me a calm feeling", "while playing the games, I was totally immersed in them", "while playing the games, it felt like I could not stop playing". Children responded on a 5-point Likert scale. Reliability statistics indicated that one item was not consistent with other measured items. As such, the final engagement scale was based on 5 items with higher scores indicating greater engagement ($M = 2.84$, $SD = 0.87$, $\alpha = .67$).

4.4.3. Creativity

Creativity was measured by the Torrance Test of Creative Thinking (TTCT; Torrance, 1988), a standardized measure of children's creativity. The TTCT consists of both a verbal and figural component, each with six different activities. Importantly, the TTCT has been used by scholars in a range of manners across the literature – with some scholars choosing to present one standardized score across the entirety of the TTCT and other using one or more subtask for assessment and analysis. To prevent participant fatigue, we opted to select and score only one task from the TTCT that has been commonly used in empirical research on children's creativity in middle childhood (e.g., Dziedziewicz, Gajda, & Karwowski, 2014): the circles task.

The circles task is a timed figural task in which children are shown a piece of paper with 36 circles. Children are allotted 10 min to draw at least one original drawing with at least one of the circles. Drawings are then coded following procedures described in the TTCT manual (Torrance, 1998). In practice, this meant that each drawing was coded for fluency and originality. Fluency reflects the number of drawings that make relevant use of at least one circle ($M = 5.37$, $SD = 5.48$, Range = 0 to 36). Originality reflects the number of original and unusual ideas children expressed in the drawings that were scored on fluency ($M = .90$, $SD = 1.60$, Range = 0 to 9). Bonus originality points are given for drawings that combine at least two circles, where more points are awarded if more circles are used to express one idea ($M = .80$, $SD = 1.42$, Range = 0 to 5). Importantly, while the entire TTCT-figural task also provides additional scoring options (i.e., "resistance to premature closure"; "elaboration"; "abstractness of titles", and "creative strengths"), the use of the circles task alone does not yield this additional information. See Fig. 2 for a scoring example.

To ensure the coding process was reliable, 10% of the drawings were scored by a second coder (Krippendorff's $\alpha = .86$). Consistent with other applications of this measure, fluency and originality were positively correlated ($r = .23$). Given that hypotheses did not distinguish between fluency and originality, a creativity score was derived by averaging together the two standardized scores (fluency and originality) with higher scores indicating greater creativity ($M = 0.00$, $SD = .79$, Range = –0.87 to 2.80). To aid future scholarship, analyses separated by fluency and originality using raw scores are also presented in the footnotes of this manuscript.

4.5. Procedure

Data collection occurred over 8 days during an annual children's film festival in the Netherlands. Data collection was conducted by two trained researchers. At the beginning of the session, children were asked several warm-up questions (e.g., "what is your favorite app?") to ensure they were comfortable with the experimenter and the setting. Then, all children completed the measure of fantastical thinking. Following this, children played their randomly assigned apps. Each app could be played for a maximum of 10 min. If the

Fig. 2. Creativity Task: Sample Responses. In the left image, participant 1031 drew the 'six eyed monster'. As only one idea was expressed, 1 point for fluency was allocated. The drawing did not appear in the originality list, resulting in one point for originality. Finally, as the child used 6 circles to express this one idea, 3 bonus points were allocated for originality. Total score = 5 points. In the right image, participant 1051 drew six drawings. Five of these made relevant use of the circles, one did not. Thus, 5 points were allocated for fluency. One drawing did not appear in the originality list, resulting in one point for originality. The participant did not combine any circles, thus no bonus points were assigned. Total score = 6 points.

Table 2
Spearman's rho correlations.

	Condition	Gender	Fantasy	Engagement	Creativity	Age	Appeal
Condition		.008	−.056	.206*	.181	.038	.411*
Gender			.147	.036	.077	.039	.118
Fantasy				.162	−.187	−.313*	.232*
Engagement					.079	−.121	.518*
Creativity						.291*	.042
Age							−.241*
Appeal							

* = Significant at <0.05 level.
** = Significant at <0.10 level.

child opted to stop play sooner than 10 min, s/he was free to do so otherwise, at the 10 min mark, the child's app play was stopped by the data collector. To avoid ordering effects, the order of apps was randomly selected for each participant. On average, children played the first app for 7.58 min (SD = 2.74) with 42% of the sample playing for the entire 10 min. On average, children played the second app for 7.32 min (SD = 2.87) with 41% of the sample playing for the entire 10 min. The correlation between app play length between the first and second app was .32 (p < .05). There was no statistically significant difference in time spent playing between conditions, although means indicates that children in the moderately discrepant app condition played with the apps for a slightly longer period of time (M = 7.81 min, SD = 1.94) than the highly discrepant condition (M = 7.12, SD = 2.52; F (1, 93) = 2.2, p = .114). During app play, children were videotaped to allow for potential video coding and qualitative analysis. Once app play was complete, children completed the engagement assessments. Finally, children completed the creativity assessment and answered several questions about app appeal. Upon conclusion, children were thanked and compensated with a small gift.

5. Results

Statistical analyses were conducted using SPSS (version 24). Analysis of Variance (ANOVA) was used to evaluate hypothesis 1 – the relationship between condition and creativity. To test hypotheses 2, 3, 4, the Hayes PROCESS macro was used. The Hayes PROCESS macro (v3) is a regression-based approach that allows for tests of statistical interaction and mediation (Hayes, 2017). To address concerns associated with residual normality, models employed bootstrapping. See Table 2 for correlation matrix of all model variables.

Hypothesis 1. Direct effect of apps on creativity

The first hypothesis predicted that moderately discrepant creative apps (i.e. apps that are appropriate for the cognitive level of 8- to 10-year-olds) would increase creativity to a greater extent than highly discrepant creative apps (i.e., apps that are inappropriate for the cognitive level of 8- to 10-year-olds). Although means (see Table 3) were in the expected direction with children in the moderately discrepant apps condition scoring higher on creativity ($M = .10$, $SD = .76$) than children that played the highly discrepant apps ($M = -0.09$, $SD = .80$), this difference was not statistically significant ($F\ (1,\ 92) = 1.27$, $p = .26$).[3] Hypothesis 1 was not supported. That said, despite this absence of a main effect, indirect effects are possible even in the absence of direct effects (Hayes, 2009). As such, investigating both mediation (H2) and moderated-mediation (H3, H4) remain testable hypotheses.

Hypothesis 2. Engagement as mediator

The second hypothesis predicted that engagement mediated the relationship between app condition and creativity, such that children that played with the moderately discrepant apps would experience increased engagement and subsequently increased creativity. Analyses revealed a marginally significant relationship between condition and engagement ($b = .34$, $SE = 0.18$, $p = .057$, 95% CI [-0.01, 0.70]) such that, as expected, children in the moderately discrepant app condition ($M = 3.02$, $SD = 0.88$) were more engaged than children in the highly discrepant app condition ($M = 2.67$, $SD = 0.83$). However, engagement did not mediate the relationship between app condition and creativity, ($b = .03$, $SE = 0.05$, 95% CI [-0.05, 0.13].[4] Hypothesis 2 was not supported.

Hypothesis 3. Gender as moderator

In addition to testing whether engagement mediated the relationship between app condition and creativity, hypothesis 3 posited that this mediation would be moderated by gender such that girls would be more engaged while playing moderately discrepant apps and therefore experience greater creativity benefits. A moderated-mediation model was conducted to test this hypothesis. Results show that gender did not moderate the influence of condition on engagement ($b = -0.45$, $SE = .36$, $p = .216$, 95% CI [-1.18, 0.27]) and subsequent creativity ($b = -0.03$, $SE = .07$, 95% CI [-0.20, 0.08]).[5] Hypothesis 3 was not supported.

Hypothesis 4. Fantasy as moderator

Just as gender was hypothesized to act as a moderator, fantastical thinking was also hypothesized to influence the relationship between app condition, engagement, and creativity such that increased fantastical thinking would lead to greater engagement with moderately discrepant apps and subsequently greater creativity. A moderated-mediation model was conducted to test this hypothesis. Results indicate that fantastical thinking did not moderate the influence of condition on engagement ($b = -.19$, $SE = .54$, $p = .724$, 95% CI [-1.27, 0.88]) and subsequent creativity ($b = -0.01$, $SE = .08$, 95% CI [-0.19, ,13]).[6] Hypothesis 4 was not supported.

5.1. Post hoc investigation

Given the reasonably robust theoretical argumentation for the proposed moderated-mediation model, the lack of statistically significant findings for this study was somewhat surprising. In the analyses, we focused on learning – namely, creativity. However, creativity is a skill that requires a significant amount of time to develop and the expectation that playing two apps would

[3] This analysis was replicated with originality and fluency separately. Both models were non-significant (*fluency:* $F\ (1,\ 92) = .28$ $p = .60$; *originality:* $F\ (1,\ 92) = 1.55$, $p = .22$), and as with the average measure, means were in the expected direction with the treatment group (*fluency:* M = 5.68, SD = 4.83; *originality:* M = 2.00, SD = 2.26) outperforming the control group (*fluency:* M = 5.08, SD = 6.05; *originality:* M = 1.42, SD = 2.17) for each attribute.

[4] This analysis was replicated with originality and fluency separately. Interpretations were identical to the average score such that engagement did not mediate the relationship between app condition and originality, ($b = .03$, $SE = 0.12$, 95% CI [-0.22, 0.31] nor app condition and fluency, ($b = .21$, $SE = .26$, 95% CI [-0.21, 0.80].

[5] This analysis was replicated with originality and fluency separately. Interpretations were identical to the average score - gender did not moderate the influence of condition, via engagement, on subsequent originality ($b = -0.03$, $SE = .18$, 95% CI [-0.51, 0.30]) or fluency ($b = -0.28$, $SE = .40$, 95% CI [-1.19, 0.48]).

[6] This analysis was replicated with originality and fluency separately. Interpretations were identical to the average score – fantastical thinking did not moderate the influence of condition, via engagement, on subsequent originality ($b = -0.01$, $SE = .19$, 95% CI [-0.44, 0.40]), or fluency, ($b = -0.12$, $SE = .50$, 95% CI [-1.27, 0.79]).

Table 3
Descriptives, Overall and by Condition.

	Overall			Highly Discrepant			Moderately Discrepant		
	Min.	Max.	*M*(*SD*)	Min.	Max.	*M*(*SD*)	Min.	Max.	*M*(*SD*)
Fantasy	1.09	2.45	1.64 (0.35)	1.09	2.45	1.67 (0.40)	1.18	2.27	1.60 (0.29)
Engagement	1.20	5	2.84 (0.87)	1.20	4.40	2.67 (0.83)	2.27	5	3.02 (0.88)
Creativity	-.87	2.80	0.00 (0.78)	-.87	2.80	-.09 (.80)	-.87	1.93	.10 (.76)
Appeal	1	5	3.55 (0.76)	1	5	3.27 (0.77)	2	5	3.85 (0.64)

Note. Creativity represents an average of the standardized scores for fluency and originality.

significantly bolster creativity, particularly using a standardized measure of creativity, may have been too high of an expectation. Indeed, other research with other types of media has indicated that such generalized gains are typically only detected in longitudinal studies with increased dosage whereas smaller near-transfer effects (i.e. content-specific gains) are more likely to be seen in experimental settings (Crawley, Anderson, Wilder, Williams, & Santomero, 1999). More recently, work with digital technology has gone so far as to suggest that interactive technology may be best (or solely) suited to near transfer learning as opposed to more generalized gains (often referred to as far transfer; see Aladé, Lauricella, Beaudoin-Ryan, & Wartella, 2016; Schroeder & Kirkorian, 2016).

That said, recall that the theoretical argumentation underlying this research (based on the MDH and flow theory) was that moderately discrepant content would be more likely to induce engagement, that this engagement would be experienced as enjoyable, and as a result, that this experience would lead to greater learning (in this case, increased creativity; Sherry, 2004). Although no game-specific (near-transfer) creativity measures were included in this study, measures of appeal *were* included for use with a different set of hypotheses. As such, we have the opportunity to assess post hoc whether moderately discrepant apps may influence the more proximal measure of appeal as opposed to the more distal measure of creativity.

Adapted from existing measures of appeal (de Droog, Buijzen, & Valkenburg, 2012; de Droog, Valkenburg, & Buijzen, 2010), app appeal was measured with two questions which asked the participant the extent to which s/he liked the first and the second app. Measured with a 5-point Likert smiley-face scale, these two scores were averaged with higher scores indicating higher appeal ($M = 3.55$, $SD = 0.76$). Using this data, we asked whether condition influenced app appeal and if this relationship was mediated by engagement. Results indicated a significant direct relationship between app condition and app appeal with moderately discrepant apps rated as significantly more appealing ($M = 3.86$, $SD = 0.64$) than highly discrepant apps ($M = 3.27$, $SD = 0.77$; $b = .46$, $SE = .14$, $p = .001$, 95% CI [0.19, 0.72]). Further, analyses indicated that engagement significantly mediated this relationship ($b = .13$, $SE = .07$, 95% CI [0.004, 0.28]) such that moderately discrepant apps were significantly more engaging than highly discrepant apps, and this increased engagement predicted greater appeal.[7] Lastly, we asked whether gender or fantasy moderated this mediation. Results indicate that neither fantasy nor gender moderated this relationship (fantasy moderation: $b = -0.07$, $SE = 0.20$, 95% CI [-0.48, 0.31]; gender moderation: $b = -16$, $SE = 0.14$, 95% CI [-0.42, 0.11]).

6. Discussion

Guided by the propositions of the MDH and flow theory, this experimental study was designed to evaluate whether moderately discrepant apps which emphasized creative play would support creativity to a greater extent than highly discrepant apps among children aged 8–10 years old via engagement. Moreover, guided by the DSMM, this study sought to understand how individual differences – specifically gender and fantastical thinking – may moderate the effects of this content. To our knowledge, this study represents the first study to investigate whether apps can support creativity in middle childhood, and moreover, is the first study to simultaneously incorporate MDH, flow theory, and DSMM into one conceptual model.

All told, results offered mixed findings. We found no support for our study hypotheses. Playing moderately discrepant creative apps did not result in increased creativity as compared to children who played highly discrepant creative apps and engagement was not a mediator in this process as predicted by MDH and flow theory. Moreover, neither gender nor fantastical thinking influenced engagement as proposed by DSMM. At first glance, these findings are disappointing. However, post hoc analyses suggest a more promising picture.

Results from the post hoc analysis indicate that children enjoyed the moderately discrepant apps more than highly discrepant apps, and further, that this increased appeal can be explained by increased engagement. This is in line with the expectations of both MDH and flow theory – namely, increased engagement should lead to increased enjoyment - which should *ultimately* support learning. As such, the lack of learning effects in this study may less reflect the inability of the apps to support learning but instead reflect methodological error. Specifically, it may be that learning effects – particularly generalized creativity gains – require more than one short experimental testing session to be moved and instead require longer sustained play and multiple moments of play (i.e. longitudinal study). Evidence from learning effects with traditional media (e.g., television; for a review see Fisch, 2004) and new media (e.g., video games; Gentile et al., 2009) certainly indicate that greater time is needed before generalized gains can be statistically detected. In fact, recent research with educational television suggests that approximately 250 min of exposure is needed for short-term learning (i.e., content that maps closely to the media) and 625 min is needed for long-term learning (i.e., generalized learning that goes beyond the media) (Linebarger, 2015). While television is certainly different than apps, when one considers the amount of app exposure (20 min) in the context of television research, duration may be a reasonable explanation for the findings presented here. It would be worthwhile to replicate this model within a longitudinal framework in which exposure duration is significantly bolstered.

At the same time, it may alternatively be the case the case that it

[7] For a complete test, we also explored the potential of a serial mediation model whereby condition predicted engagement, which subsequently predicted appeal, which subsequently predicted creativity. Analyses showed no support for serial mediation in that, while condition did predict engagement and while engagement did predict appeal (as reported), appeal did not predict creativity (or either of its domains - *total indirect effect on creativity:* $b = .02$, $SE = .07$, 95% CI [-0.10, 0.18]; *total indirect effect on fluency:* $b = .50$, $SE = .44$, 95% CI [-23, 1.48]; *total indirect effect on originality:* $b = -0.10$, $SE = .24$, 95% CI [-0.55, 0.40]).

is not the length of duration per se that is particularly problematic but the type of outcome measure. In this study, we used a far transfer task (see, for example, Aladé et al., 2016) – where the task involves transferring the learned knowledge to a new situation – by measuring creativity with a standardized measure. This is in contrast with a near transfer task (in which the task is most closely aligned with the content of the learning experience). Far transfer is generally considered a more complex skill that occurs later in the learning process. Since each of the apps had its own approach to building in creativity, we felt that a broader task which looked at creativity without app specificity would (1) allow us to use a standardized measure as opposed to a set of researcher-developed measures (improving the measurement validity of the study) that would facilitate comparison across conditions, and (2) would minimize potential issues of fatigue as the selected measure is easy-to-administer, relatively quick, and interpreted positively by young children. However, in retrospect, this may mean we missed more subtle knowledge gains that a near-transfer task could have detected. As noted above, scholars have recently gone so far as to suggest that the interactivity inherent in technology such as apps may be most helpful in contexts that are highly similar to the original learning task (i.e., near transfer) but lead to little lasting effects once the transfer task is far removed from the original learning context (i.e., far transfer tasks; Aladé et al., 2016; Schroeder & Kirkorian, 2016). As such, it may be the case that learning did occur from these apps – but that the measurements used in this study were too far removed from the app content. Replicating this study with *both* near and far transfer measurement points, along with increased exposure duration to the apps, would be a justified next step.

In addition to the main and mediation effects investigated in this study, two individual differences variables – gender and fantastical thinking – were also investigated. The results indicate that neither variable interacted with condition to influence engagement. This is not necessarily contrary to the propositions of the DSMM. The DSMM argues that media processing will be moderated by individual differences, but not necessarily moderated by all individual difference variables. In the study, two variables frequently discussed in the creativity literature were evaluated – but there are, of course, many others that would make sense in future research. For example, research suggests that app engagement might be influenced by curiosity (Arnone, Small, Chauncey, & McKenna, 2011), intelligence (Kershner & Ledger, 1985), and degree of self-expression (Runco, 2006). All of these variables are reasonable points for additional research inquiry.

Moreover, in addition to future research with other individual differences, it is also important to think about the lack of moderation effects found in this study. For example, with gender, it is somewhat reassuring to see that both boys and girls engaged with the app content in similar ways. A look at the media and toy landscape for youth today often highlights the "pink and blue" phenomenon with content explicitly demarcated as appropriate for girls or boys. Here, the results show that when content does not try to target a specific gender (an inclusion criterion of this study), both genders can share similar experiences. Rather than creating "pink" or "blue" app experiences, these findings suggest that media creators should focus more on the quality of their app since a high quality app can be equally engaging, and appealing, for both boys and girls.

As to fantastical thinking, the lack of significant moderation may be due to measurement error. Strangely, although not significant, data patterns suggested that children with a tendency to engage in greater fantastical thinking were less engaged with the app content. This is the opposite of what was expected. In reflecting on the short form of the scale used here, the items tend to reflect the "daydreaming" component of fantasy more than anything else (i.e., aspects of imagination and pretending are more limited) which certainly may lead to a tendency to be less engaged in specific situations. As such, the content validity of the scale may be less assured than originally thought. Efforts to replicate this study with an improved measure would be reasonable.

6.1. Implications

While many of the original expectations guiding this study were not supported, this study still offers important theoretical, methodological, and practical implications. Theoretically, the findings from the post hoc investigation on appeal provide increased confidence in the integration of MDH, flow theory, and DSMM. This is the first time, to our knowledge, that these three concepts have been integrated into one model. We see this conceptual model as an important steppingstone for future scholarship on how children select, experience, and are affected by digital media content. We hope that this theoretical convergence serves as inspiration to other scholars interested in understanding boundary conditions which explain for which children effects are present, how these effects unfold, and what these effects might look like.

Methodologically, the lack of significant effects of app characteristics on performance in the creativity task – in conjunction with the significant findings for appeal – suggest that a longitudinal component with more precise measurements (e.g., near transfer tasks) is a necessary next step towards understanding app effectiveness. A short-term experiment with generalized outcome measurement, such as this one, may be insensitive to the potentially lengthier and/or more nuanced process associated with children's learning from apps. Finally, from a practical perspective, this study provides initial evidence to indicate that successful media content will be that content which is congruent with the target group's cognitive, social, and emotional development. Content that is too different from a user's development may be interpreted as too easy or too difficult, and as such, is unlikely to result in engagement, appeal, or subsequent learning.

6.2. Looking forward

In addition to replicating this study with a longitudinal component in order to increase exposure and be more sensitive to the potential time required for generalized gains in creativity, there are also other areas where this study can be improved and expanded in future research. First, although an attempt was made to collect detailed information about the child's home environment for additional potential moderators and/or control variables, the response rate for the parent survey was quite low. The study was conducted during a large festival and, as such, many parents preferred to view other festival activities or engage with their other children rather than complete the survey. In future work, it would be valuable to consider whether the home environment may influence how children are affected. For example, research has shown that children who have greater technological familiarity tend to be less engaged with tablet-based media than their peers who have less technological skills (Krcmar & Cingel, 2014), pointing to potential novelty effects (Henderson & Yeow, 2012).

Second, this study used previously existing apps as opposed to apps designed specifically for this study in question. While the use of existing apps was deliberately chosen to enhance the external validity of the study – this enhancement is done as the loss of internal validity. Had we opted to use apps created specifically for this study – we would have had much stronger control of the study manipulation, ensuring equivalence on all domains outside of 'discrepancy level', which may have led to more pronounced or

otherwise different effects. In work with video games, for example, scholars have shown differences in the strength of relationships based on whether the stimuli is pre-existing or experimenter-developed (Powers et al., 2013). Furthermore, while we attempted to alleviate some limitations of using existing apps by testing more than 1 app per condition (e.g., to avoid potential app-specific effects), in doing so, we may also have introduced unsystematic variance into our analyses. This tension between external and internal validity is certainly a challenge in this study, and one that would welcome replication and extension. In particular, while budgetary limitations prevented the development of an experimental-developed app in this study, future work with such resources would be a welcomed addition to the field.

Third, although we have argued that there is a need for both near and far transfer measures in future scholarship, we would also suggest that a more comprehensive measure of creativity be implemented. The limited time that we had with each child in our sample meant that administering the TTCT in its entirety was not a feasible option. As such, we opted to select one subtask to measure creativity, treating creativity as unidimensional (as the TTCT originally was designed). However, in more recent years, scholars have challenged this unidimensional perspective – suggesting instead that creativity may be better thought of as a multidimensional construct consisting of both innovative and adaptive orientations (Kim, 2006). In this perspective, innovative creative thinking refers to thinking that results in new solutions whereby adaptive creative thinking refers to thinking in which existing solutions are applied to new scenarios (Kim & Pierce, 2013, pp. 35–40). An innovative orientation is said to be captured in quick and novel responses whereas an adaptive orientation is reflected through thoughtful and elaborate responses (Kim, 2006). While Kim (2006) has shown that the entire TTCT does a reasonable job of capturing both orientations, the subtest employed here would be considered to represent the innovative domain (Kim, 2006; Krumm, Filippetti, Lemos, Koval, & Balabanian, 2016). As such, by replicating this study with the entire TTCT, scholars would be able to investigate this multidimensional perspective in more detail. For example, it may be that apps are better suited to supporting the adaptive side of creativity – at least for some children.

Within the context of measurement, it is also crucial to recognize that measuring engagement was a significant challenge in this study. Although not reported in the measurement section of this manuscript, this study actually employed three measures of engagement. During the planning stage of this study, it became clear that there were conflicting ideas of how to measure young children's engagement. In an effort to be comprehensive as well as potentially triangulate this complex concept, we opted to include three measurement approaches. In practice, this meant that we also used a rating scale from Calvert, Strong, Jacobs, and Conger (2007) in which engagement (operationalized as physical and verbal engagement) was assessed during the iPad game play. Specifically, during game play, the researcher recorded the level of engagement on thirty-second intervals using a 4-point scale ranging from 'no engagement' (0) to 'enthusiastic engagement' (3). Previously, this scale has only been used with young children watching television. In this study, not only did the scale achieve low variability, but both study data collectors noted that the measurement scale seemed to measure distraction from the apps as opposed to engagement with them. For example, in our review of the video data in comparison with the engagement codes, it seemed that coders were most likely to record "enthusiastic engagement" when the child was dancing/moving/stopping app play in some way unrelated to the app – thus rather than tapping into engagement, it seemed to almost tap into the reverse. Indeed, in analyses, we found that this measure correlated in the reverse direction (−.17, NS) with children's self-reported engagement. Given the video data, we felt that we could not argue that the measure was a valid measure of engagement and opted to exclude it from further consideration.

In addition to the Calvert et al. (2007) scale, all data collection sessions were video-taped in order to apply another engagement measurement – this time, an observational protocol developed by Roskos, Burstein, Shang, and Gray (2014) that was originally used to assess engagement with eBooks. All videos were coded for thirteen different behaviors that were argued by Roskos and colleagues to reflect children's engagement (defined as a combination of attention, interest, and enjoyment where children look, see, and listen with apparent pleasure). In this scale, engagement included binary measurement of concepts such as listening, looking, touching, moving, gesturing, making noises, and facial expression. After coding all videos (and achieving intercoder reliability), we found that there was almost no variance in engagement across the entirety of the sample or by condition. In retrospect, we believe the binary option of the scale was too limiting since all children were marked as having high engagement. Perhaps with e-books in which the story is read to them (as was originally used), this binary approach is reasonable. But with apps that call upon children's interaction as part of their use, it would be nearly impossible to advance in the app without all of these behaviors. Given this lack of measurement validity, similar to the Calvert scale, we felt it was inappropriate to include this measure in our analyses.

As a result, despite three measurement approaches, it was only the adapted self-report measure of engagement that yielded variance across the study. Although this self-report measure, in terms of the post hoc analyses, did yield theoretically consistent findings, future efforts to identify the best approach to measure engagement seems crucial given the importance of engagement in digital media play. In particular, it would be interesting to extend this self-report measure in such a way that it independently reflects the cognitive, emotional, and physiological aspects of engagement since it may be the case that – depending on the digital media content or format – different attributes of engagement may be more affected more than others. Moreover, it would be interesting to identify whether a valid observational protocol (similar to the ones tested here) could be developed to measure engagement, particularly for use with younger audiences, since the self-report of engagement requires a level of meta-cognitive awareness that youngers users may not have sufficiently developed yet.

Fifth, also related to measurement, there are few good self-report measures available for measuring individual differences in children – and those that do exist tend to be lengthy and likely to result in participant fatigue when used in conjunction with other measures. The fantasy measure, for example, was originally 45 items. Using the available psychometric properties of this variable, we shortened this scale to a more manageable length of 11 items. However, in doing so, it is likely that the content validity of the measure was decreased – which may explain the lack of findings for this measure. Efforts to identify ways to measure individual differences in non-obtrusive ways among children are important both for replication and extension of this study as well as for youth and media scholarship more generally.

Finally, this study represented a small convenience sample of Dutch children aged 8–10 years old. The relatively small size of the sample may have resulted in underpowered analyses – particularly for the moderated mediation models. Further, the convenience recruitment means that it is unclear whether findings are generalizable to other children of this age. Efforts to replicate this work with a larger and more representative sample of Dutch children would provide increased statistical power for more nuanced analyses as well as an indication to the generalizability (or specificity) of study findings.

7. Conclusion

The children's app market is continuing to grow in size and scope, and with it, children's time spent with apps continues to increase. As such, there is no doubt that there will be increased interest from parents, caregivers, educators, and practitioners as to the opportunities and consequences of children's app use as well as best practices in app design (Hirsh-Pasek et al., 2015). With only a handful of existing studies currently investigating any aspect of children's apps, this study offers an important first look as to how creative apps may affect children in middle childhood. Although findings suggest that short-term creative app use does not influence creativity skills, findings do indicate that apps which are tailored to the user's developmental lead to increased engagement and subsequent appeal. Considering the importance of appeal for the continued use of media content (Hamlen, 2011), this finding offers an important contribution to the growing body of work on children's apps. Educational app developers would be prudent to thoughtfully consider the target age of their content early on, and to work to ensure that their content is neither too easy nor too difficult for their user. In other words, they should work to ensure that their content is moderately discrepant from their users' skills and knowledge. In doing so, they increase the likelihood that their content reaches its desired goals, and simultaneously, increase the likelihood that the time children spend with their apps is time well spent.

Acknowledgements

This work was supported, in part, by the Graduate School of Communication and the Amsterdam School of Communication Research ASCoR at the University of Amsterdam as well by the Cinekid organization. These organizations were unassociated with the study design, analyses, or reporting of this work. Portions of this work were presented at the 2016 annual meeting of the International Communication Association in San Juan, Puerto Rico.

References

Aladé, F., Lauricella, A. R., Beaudoin-Ryan, L., & Wartella, E. (2016). Measuring with Murray: Touchscreen technology and preschoolers' STEM learning. *Computers in Human Behavior, 62*, 433–441. https://doi.org/10.1016/j.chb.2016.03.080.

Amabile, T. (2012). *Componential theory of creativity*. Boston: Harvard Business School.

Annetta, L. A., Minogue, J., Holmes, S. Y., & Cheng, M.-T. (2009). Investigating the impact of video games on high school students' engagement and learning about genetics. *Computers & Education, 53*(1), 74–85. https://doi.org/10.1016/j.compedu.2008.12.020.

Arnone, M. P., Small, R. V., Chauncey, S. A., & McKenna, H. P. (2011). Curiosity, interest and engagement in technology-pervasive learning environments: A new research agenda. *Educational Technology Research & Development, 59*(2), 181–198. https://doi.org/10.1007/s11423-011-9190-9.

Aziz, N. A. A. (2013). Children's interaction with tablet applications: Gestures and interface design. *International Journal of Computer and Information Technology, 2*(3), 447–450.

Bandura, A. (1977). *Social learning theory*. New York: General Learning Press.

Brockmyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M., & Pidruzny, J. N. (2009). The development of the game engagement questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology, 45*(4), 624–634. https://doi.org/10.1016/j.jesp.2009.02.016.

Broekman, F. L., Piotrowski, J. T., Beentjes, H. W., & Valkenburg, P. M. (2016). A parental perspective on apps for young children. *Computers in Human Behavior, 63*, 142–151.

Burnard, P., & Younker, B. A. (2004). Problem-solving and creativity: Insights from students' individual composing pathways. *International Journal of Music Education, 22*(1), 59–76.

Cacioppo, J. T., Pett, R. E., & Stoltenberg, C. D. (1985). Processes of social influence: The Elaboration Likelihood model of persuasion. In P. C. Kendall (Ed.), *Advances in cognitive-behavioral research and therapy. Advances in cognitive-behavioral research and therapy* (Vol. 4, pp. 215–274). San Diego, CA: Academic Press.

Calvert, S. L., Strong, B. L., Jacobs, E. L., & Conger, E. E. (2007). Interaction and participation for young Hispanic and Caucasian girls' and boys' learning of media content. *Media Psychology, 9*, 431–445. https://doi.org/10.1080/15213260701291379.

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*. Retrieved from New York http://www-tc.pbskids.org/read/files/cooney_learning_apps.pdf.

Cohen, M., Hadley, M., & Frank, M. (2011). *Young children, apps & iPad*. New York, NY, USA: Michael Cohen Group LLC.

Crawley, A. M., Anderson, D. R., Wilder, A., Williams, M., & Santomero, A. (1999). Effects of repeated exposure to a single episode of the television program *Blue's Clues* on the viewing behaviors and comprehension of preschool children. *Journal of Educational Psychology, 91*(4), 630–637. https://doi.org/10.1037/0022-0663.91.4.630.

Csikszentmihalyi, M. (1997). *Finding flow: The psychology of engagement with everyday life*. New York: Basic Books.

Dansky, J. L. (1980). Make-believe: A mediator of the relationship between play and associative fluency. *Child Development, 51*, 576–579. https://doi.org/10.2307/1129296.

Dickey, M. D. (2005). Engaging by design: How engagement strategies in popular computer and video games can inform instructional design. *Educational Technology Research & Development, 53*(2), 67–83. https://doi.org/10.1007/BF02504866.

Dickson, J. (2014). *Educational iPad apps still popular with nine to 11s*. Kidscreen. Retrieved from http://kidscreen.com/2014/09/11/educational-ipad-apps-still-popular-with-nine-to-elevens/.

de Droog, S. M., Buijzen, M., & Valkenburg, P. M. (2012). Use a rabbit or a rhino to sell a carrot? The effect of character-product congruence on children's liking of healthy foods. *Journal of Health Communication, 17*(9), 1068–1080. https://doi.org/10.1080/10810730.2011.650833.

de Droog, S. M., Valkenburg, P. M., & Buijzen, M. (2010). Using brand characters to promote young children's liking of and purchase requests for fruit. *Journal of Health Communication: International Perspectives, 16*(1), 79–89. https://doi.org/10.1080/10810730.2010.529487.

Dziedziewicz, D., Gajda, A., & Karwowski, M. (2014). Developing children's intercultural competence and creativity. *Thinking Skills and Creativity, 13*, 32–42. https://doi.org/10.1016/j.tsc.2014.02.006.

Falloon, G. (2013). Young students using iPads: App design and content influences on their learning pathways. *Computers & Education, 68*, 505–521. https://doi.org/10.1016/j.compedu.2013.06.006.

Fisch, S. M. (2004). *Children's learning from educational television*. Mahwah, NJ: Lawrence Erlbaum.

Flynn, J. R. (1999). Searching for justice: The discovery of IQ gains over time. *American Psychologist, 54*(1), 5. https://doi.org/10.1037/0003-066X.54.1.5.

Garzotto, F. (2007). Investigating the educational effectiveness of multiplayer online games for children. In *Paper presented at the Proceedings of the 6th international conference on Interaction design and children*.

Gentile, D. A., Anderson, C. A., Yukawa, S., Ihori, N., Saleem, M., Lim Kam, M., & Sakamoto, A. (2009). The effects of prosocial video games on prosocial behaviors: International evidence from correlational, longitudinal, and experimental studies. *Personality and Social Psychology Bulletin, 35*(6), 752–763. https://doi.org/10.1177/0146167209333045.

Hamlen, K. R. (2011). Children's choices and strategies in video games. *Computers in Human Behavior, 27*(1), 532–539. https://doi.org/10.1016/j.chb.2010.10.001.

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs, 76*(4), 408–420. https://doi.org/10.1080/03637750903310360.

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Publications.

Henderson, S., & Yeow, J. (2012). iPad in education: A case study of iPad adoption and use in a primary school. In *Paper presented at the System science (HICSS), 2012 45th Hawaii international conference*. IEEE.

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. https://doi.org/10.1177/1529100615569721.

Holloway, D., Green, L., & Livingstone, S. (2013). *Zero to eight: Young children and their internet use*. Retrieved from London, UK http://eprints.lse.ac.uk/52630/1/Zero_to_eight.pdf.

Howard-Jones, P. A., & Demetriou, S. (2009). Uncertainty and engagement with learning games. *Instructional Science, 37*(6), 519. https://doi.org/10.1007/s11251-008-9073-6.

Huizenga, J., Admiraal, W., Akkerman, S., & Dam, G. t. (2009). Mobile game-based learning in secondary education: Engagement, motivation and learning in a mobile city game. *Journal of Computer Assisted Learning, 25*(4), 332–344. https://doi.org/10.1111/j.1365-2729.2009.00316.x.

Inal, Y., & Cagiltay, K. (2007). Flow experiences of children in an interactive social game environment. *British Journal of Educational Technology, 38*(3), 455–464. https://doi.org/10.1111/j.1467-8535.2007.00709.x.

Jackson, L. A., Witt, E. A., Games, A. I., Fitzgerald, H. E., von Eye, A., & Zhao, Y. (2012). Information technology use and creativity: Findings from the children and technology Project. *Computers in Human Behavior, 28*(2), 370–376. https://doi.org/10.1016/j.chb.2011.10.006.

Kabali, H. K., Irigoyen, M. M., Nunez-Davis, R., Budacki, J. G., Mohanty, S. H., Leister, K. P., et al. (2015). Exposure and use of mobile media devices by young children. *Pediatrics, 136*(6), 1–9. https://doi.org/10.1542/peds.2015-2151.

Kershner, J. R., & Ledger, G. (1985). Effect of sex, intelligence, and style of thinking on creativity: A comparison of gifted and average IQ children. *Journal of*

Personality and Social Psychology, 48(4), 1033. https://doi.org/10.1037/0022-3514.48.4.1033.
Kim, K. H. (2006). Is creativity unidimensional or multidimensional? Analyses of the torrance tests of creative thinking. *Creativity Research Journal, 18*(3), 251–259. https://doi.org/10.1207/s15326934crj1803_2.
Kim, K. H., & Pierce, R. A. (2013). *Adaptive creativity and innovative creativity. In encyclopedia of creativity, invention, innovation and entrepreneurship*. New York: Springer.
Krcmar, M. (2010). Can social meaningfulness and repeat exposure help infants and toddlers overcome the video deficit? *Media Psychology, 13*(1), 31–53. https://doi.org/10.1080/15213260903562917.
Krcmar, M., & Cingel, D. P. (2014). Parent–child joint reading in traditional and electronic formats. *Media Psychology, 17*(3), 262–281. https://doi.org/10.1080/15213269.2013.840243.
Krumm, G., Filippetti, V. A., Lemos, V., Koval, J., & Balabanian, C. (2016). Construct validity and factorial invariance across sex of the torrance test of creative thinking–figural form a in Spanish-speaking children. *Thinking Skills and Creativity, 22*, 180–189. https://doi.org/10.1016/j.tsc.2016.10.003.
Linebarger, D. L. (2015). Super why! to the rescue: Can preschoolers learn early literacy skills from educational television. *International Journal for Cross-Disciplinary Subjects in Education, 6*(1), 2060–2068. https://doi.org/10.20533/ijcdse.2042.6364.2015.0286.
Maytom, T. (2014). *Majority of UK Under-45s now own tablet devices*. Retrieved from http://mobilemarketingmagazine.com/majority-of-uk-under-45s-now-own-tablet-devices-new-study-reveals.
McPake, J., Plowman, L., & Stephen, C. (2013). Pre-school children creating and communicating with digital technologies in the home. *British Journal of Educational Technology, 44*(3), 421–431. https://doi.org/10.1111/j.1467-8535.2012.01323.x.
Mullineaux, P. Y., & Dilalla, L. F. (2009). Preschool pretend play behaviors and early adolescent creativity. *Journal of Creative Behavior, 43*(1), 41–57. https://doi.org/10.1002/j.2162-6057.2009.tb01305.x.
Mumford, M. D., & Gustafson, S. B. (1988). Creativity syndrome: Integration, application, and innovation. *Psychological Bulletin, 103*(1), 27–43. https://doi.org/10.1037//0033-2909.103.1.27.
Nusbaum, E. C., & Silvia, P. J. (2011). Are intelligence and creativity really so different?: Fluid intelligence, executive processes, and strategy use in divergent thinking. *Intelligence, 39*(1), 36–45. https://doi.org/10.1016/j.intell.2010.11.002.
O'Hare, E. (2014). *Mobile apps for children*. Retrieved from https://www.cinekid.nl/sites/cinekid.nl/files/bijlage/Research%20Report%20-%20Mobile%20Apps%20for%20Children.pdf.
Pavlas, D. (2010). *A model of flow and play in game-based learning: The impact of game characteristics, player traits, and player states*. Doctoral dissertation. Florida: University of Central Florida Orlando.
Powers, K. L., Brooks, P. J., Aldrich, N. J., Palladino, M. A., & Alfieri, L. (2013). Effects of video-game play on information processing: A meta-analytic investigation. *Psychonomic Bulletin & Review, 20*(6), 1055–1079. https://doi.org/10.3758/s13423-013-0418-z.
Prensky, M. (2001). Digital natives, digital immigrants part 1. *On The Horizon, 9*(5), 1–6.
Rideout, V. (2014). *Learning at home: Families' educational media use in America* (Retrieved from New York, NY).
Rosenfeld, E., Huesmann, L. R., Eron, L. D., & Torney-Purta, J. V. (1982). Measuring patterns of fantasy behavior in children. *Journal of Personality and Social Psychology, 42*(2), 347. https://doi.org/10.1037/0022-3514.42.2.347.
Roskos, K., Burstein, K., Shang, Y., & Gray, E. (2014). Young Children's engagement with e-books at School: Does device matter? *SAGE Open, 4*(1), 1–9. https://doi.org/10.1177/2158244013517244.
Runco, M. A. (2006). Reasoning and personal creativity. In J. C. Kaufman, & J. Baer (Eds.), *Creativity and reason in cognitive development* (pp. 99–116). New York: Cambridge University Press.
Runco, M. A., & Okuda, S. M. (1988). Problem discovery, divergent thinking, and the creative process. *Journal of Youth and Adolescence, 17*(3), 211–220. https://doi.org/10.1007/BF01538162.
Runco, M. A., & Pezdek, K. (1984). The effect of television and radio on children's creativity. *Human Communication Research, 11*(1), 109–120. https://doi.org/10.1111/j.1468-2958.1984.tb00040.x.
Russ, S. W., Robins, A. L., & Christiano, B. A. (1999). Pretend play: Longitudinal prediction of creativity and affect in fantasy in children. *Creativity Research Journal, 12*(2), 129–139. https://doi.org/10.1207/s15326934crj1202_5.
Schroeder, E. L., & Kirkorian, H. L. (2016). When seeing is better than doing: Preschoolers' transfer of STEM skills using touchscreen games. *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.01377.
Sherry, J. L. (2004). Flow and media enjoyment. *Communication Theory, 14*(4), 328–347. https://doi.org/10.1111/j.1468-2885.2004.tb00318.x.
iOS Top Apps Charts. (2015). Retrieved from https://www.appannie.com/apps/ios/top/netherlands/kids/ages-9-11.
Torrance, E. P. (1988). The nature of creativity as manifest in its testing. In R. J. Sternberg (Ed.), *The nature of creativity: Contemporary psychological perspectives* (pp. 43–75). New York: Cambridge University Press.
Torrance, E. P. (1998). *The torrance tests of creative thinking norms—technical manual figural (streamlined) forms A & B*. Bensenville, IL: Scholastic Testing Service.
Valkenburg, P. M., & Beentjes, J. W. J. (1997). Children's creative imagination in response to radio and television stories. *Journal of Communication, 47*(2), 21–38. https://doi.org/10.1111/j.1460-2466.1997.tb02704.x.
Valkenburg, P. M., & Cantor, J. (2000). Children's likes and dislikes of entertainment programs. In D. Zillman, & P. Vorderer (Eds.), *Media entertainment: The psychology of its appeal* (Vol. 11, pp. 135–152). New Jersey: Erlbaum.
Valkenburg, P. M., & Peter, J. (2006). Fantasy and imagination. In J. Bryant, & P. Vorderer (Eds.), *The psychology of entertainment* (pp. 105–117). Mahwah, NJ: Erlbaum.
Valkenburg, P. M., & Peter, J. (2013). The differential susceptibility to media effects model. *Journal of Communication, 63*(2), 221–243. https://doi.org/10.1111/jcom.12024.
Valkenburg, P. M., & Piotrowski, J. T. (2017). *Plugged in: How media attract and affect youth*. Yale University Press.
Valkenburg, P. M., & Vroone, M. (2004). Developmental changes in infants' and toddlers' attention to television entertainment. *Communication Research, 31*(3), 288–311. https://doi.org/10.1177/0093650204263435.
Vorderer, P., Klimmt, C., & Ritterfeld, U. (2004). Enjoyment: At the heart of media entertainment. *Communication Theory, 14*(4), 388–408. https://doi.org/10.1093/ct/14.4.388.
Weber, S., & Mitchell, C. (2008). *Imaging, keyboarding, and posting identities: Young people and new media technologies. Youth, identity, and digital media*. Cambridge: MIT Press.
Woolley, J. D. (1997). Thinking about fantasy: Are children fundamentally different thinkers and believers from adults? *Child Development, 68*(6), 991–1011. https://doi.org/10.2307/1132282.