What helps children learn difficult tasks: A teacher's presence may be worth more than a screen

Abstract

What helps children learn: is it a presence of a live teacher or an interaction with the learning materials? Addressing this question, we manipulated a teacher's presence (on-screen vs. present) and activity (observing vs. doing) while teaching children about the properties of geometric shapes. Five-year-olds (n=215) completed two shape-sorting tasks in which they distinguished between typical, atypical and non-valid shapes. In between these tasks, they took part in one of four training sessions: *doing teacher-present*, *observing teacher-present*, *doing teacher-on-screen* and *observing teacher-on-screen*. Although children's shape knowledge improved across all training conditions, learning showed an interaction between teacher presence and task difficulty. In a teacher's presence, children learned more about the most difficult (atypical) shapes, irrespective of activity. It may be the social interaction, associated with a teacher's presence, that enhances learning. Conversely, physically taking part in interactive touchscreen training did not result in more learning than passive screen viewing.

Introduction

In less than a decade since the introduction of the Apple iPad and Samsung Galaxy Tab (Nield, 2016), the tablet has become the only screen device, other than the television, that is consistently used at home by the majority of British 3-15-year-old children (Ofcom, 2017). Its touch-sensitive screen can be controlled with simple finger gestures making it accessible even for the youngest of children (Lovato & Waxman, 2016). This intuitive tactile interaction with on-screen content, combined with access to tens of thousands of educational applications, means that tablets have the potential to be useful educational tools, which could support young children's learning at home and in schools (Hirsh-Pasek et al., 2015). However, the enthusiasm for tablets as educational aids should be tempered by the scarcity of evidence showing their effectiveness (e.g., Lovato & Waxman, 2016; Radesky, Schumacher, & Zuckerman, 2015), particularly, in comparison to traditional person-led teaching.

Traditional passive screen media (e.g., TV, DVDs) have been used to support learning for decades, and there is well-established evidence that watching high-quality educational content can improve children's early academic skills (e.g., Baydar, Kağitçibaşi, Küntay, & Gökşen, 2008; Linebarger, Kosanic, Greenwood, & Doku, 2004; Wright et al., 2001). However, these benefits appear greatest for pre-schoolers (see Kostyrka-Allchorne, Cooper, & Simpson, 2017 for a review). Moreover, while educational television seems effective in teaching children simple content (e.g., new vocabulary; Baydar et al., 2008), it may be inadequate for teaching more cognitively challenging material (e.g., Linebarger et al., 2004). Finally, research that specifically compares learning from passive screen media to learning from 'live' interactions with an adult, consistently shows that infants and children learn most in the presence of a teacher (DeLoache et al., 2010; Krcmar, 2011; Krcmar, Grela, &

Lin, 2007; Kuhl, Tsao, & Liu, 2003; Roseberry, Hirsh-Pasek, Parish-Morris, & Golinkoff, 2009; Strouse, O'Doherty, & Troseth, 2013).

Kuhl and colleagues (2003) found that 9-month-olds have a remarkable ability to learn a non-native language during interactions with an adult repeated over a short term. However, the same level of exposure to foreign language videos did not have any benefit for learning (Kuhl et al., 2003). Similarly, while infants (< 24months) learned novel vocabulary in a single teaching session involving interactions with an adult, they did not benefit from watching a video specifically designed to teach this material (Krcmar, 2011; Krcmar et al., 2007). Moreover, recent evidence suggests that older children may also struggle to encode task-relevant information presented in a video. Reiß, Krüger, and Krist (2019) showed that 5-year-olds, who watched a live demonstration of a false belief task (either directly or through a one-way mirror) outperformed children in a video condition. Together, these findings suggest that the effectiveness of using videos as educational or instructional aids may be constrained by the lack of social interactions that are inherent to passive screen learning.

Further support for this proposal comes from studies where teaching from videos was supplemented with the help from a 'live' teacher. Roseberry et al. (2009) showed that toddlers (< 36 months) learned new verbs from a video, but only when on-screen teacher was supported with a tutorial provided by the experimenter. Older children (> 36 months) did learn something, but not as much as when supported by the experimenter. Finally, Strouse et al. (2013) showed that the benefits of adult support during learning from passive media extended to older children. Although repeated exposure to a story-like video resulted in new vocabulary learning, parent-child interactions during viewing, for example, discussing the content, further improved comprehension and general vocabulary growth. Interestingly, Strouse et al.

(2013) showed that supplementing the videos with a recording of an actress drawing children's attention to the content presented on the screen and asking questions improved story comprehension but did not improve general vocabulary. This suggests that despite increased understanding of the content, without parent-child interaction, learning new vocabulary from the videos remained difficult.

Except for the study of Reiß et al. (2019), the previous evidence is largely restricted to language learning and comes mainly from studies with infants and toddlers. Nevertheless, it clearly highlights the importance of social interaction in learning (for a detailed discussion of the social learning hypothesis see Kuhl, 2007). That passive media struggles to provide the active and engaged teaching which seems necessary to learn complex skills may be due by their lack of *interactivity*. Specifically, passive media afford few opportunities to deliver "…forms of content designed to facilitate active and creative use by young children and to encourage social engagement with other children and adults." (Radich, 2013, p.18). However, touchscreens might be well suited to address this limitation.

Touchscreens are highly responsive to users' input and can support personalised learning through contingent feedback and by adapting the level of content difficulty (Hirsh-Pasek et al., 2015). In this way, they may allow a certain degree of reciprocity akin to that present in social child-teacher interaction (Huber et al., 2016; Kuhl, 2007), and thus may improve learning beyond what can be achieved through passive media. Importantly, touchscreens bring bodily actions into the educational experience (Galetzka, 2017). A major premise of embodied cognition is that physical interaction with the environment benefits learning throughout life (Kontra, Goldin-Meadow, & Beilock, 2012). The addition of action and gesture may be particularly valuable when learning abstract subjects, such as mathematics (Alibali

& Nathan, 2012; Goldin-Meadow, Cook, & Mitchell, 2009). Past research has shown that physical interaction in math teaching (e.g., manipulating beads or counters) improved children's understanding of formal concepts such as, for example, prime numbers or addition and subtraction of fractions, more than relying on less interactive methods (for a discussion see Moyer, 2001).

Some evidence suggests interaction with a touchscreen can also improve mathematical understanding (Agostinho et al., 2015; Dubé & McEwen, 2015). Agostinho et al. (2015) demonstrated that, compared to reading from the screen, tracing on-screen information with a finger improved children's understanding of temperature graphs. Although, not all interactions with a touchscreen may be equally helpful in supporting conceptual knowledge development. Dubé and McEwen (2015) showed that on-screen dragging, but not tapping, improved adults' understanding of a mathematical concept. Moreover, there remains a concern that interacting with touchscreen is a poor substitute for hands-on manipulation of physical objects (Spitzer, 2013). Thus, it is not clear whether touchscreen learning will lead to the same level of understanding as interaction with physical objects (Miklashevsky & Fischer, 2017)

Research that compared pre-schoolers' (< 5years) learning of STEM-related (Science, Technology, Engineering, Mathematics) concepts using interactive games with passive videos presented even more complex findings. Aladé, Lauricella, Beaudoin-Ryan, and Wartella (2016) demonstrated that children who watched a video or played an interactive touchscreen game learned more than the control children, but only when tested with materials that were perceptually similar to those used during training. When test materials were perceptually different, only the children who watched a video outperformed the control group. Similarly, Schroeder and Kirkorian

(2016) demonstrated that when training and test materials were perceptually similar (i.e., both used 2D images), older pre-schoolers learned equally well by playing an interactive game or watching a video of this game. However, when the testing context was different (3D objects) to the teaching context (2D images), learning with a non-interactive video was again more effective. Thus, the findings from these two studies suggest that for children younger than five years, the benefits of learning with touchscreens are restricted by context.

The present study

Previous research has investigated the role of a teacher's presence separately from the role of interactivity. Crucially, prior studies do not answer the question of whether touchscreen interactivity can compensate for the lack of a teacher's presence. Thus, this study aimed to investigate what helped children learn. Specifically, we manipulated a teacher's presence (on-screen vs. present) and activity (observing vs. doing) in a training session, in which children were taught the formal definitions of 2D shapes (e.g., a rectangle has four straight sides). Five-year-old participants were assigned to one of four training conditions: *doing with teacher-present* (child watched a teacher and interacted with physical objects while receiving feedback); *observing, with teacher-present* (child just watched a teacher); *doing with teacher-on-screen* (child watched a video of a teacher and interacted with touchscreen objects while receiving feedback); *observing with teacher-on-screen* (child just watched a video of a teacher).

The use of 2D shapes as learning material reduced a potential confound. Teacher's presence (on-screen vs. present) also determined whether children interacted with 2D objects (teacher on-screen) or 3D objects (teacher present).

Research with infants suggest that that they learn better about 3D objects from training with 3D objects (for a discussion see Barr, 2010), and it is possible that this difficulty with transferring learning between contexts (2D to 3D) persists into later childhood (e.g., Reiß et al., 2019). Teaching children about 2D shapes minimised this contextual difference, as both the teacher on-screen and teacher present conditions presented 2D shapes.

To test children's shape learning, a sorting task (Fisher, Hirsh-Pasek, Newcombe, & Golinkoff, 2013) was administered before and after the training. By testing three shape types (typical, atypical, non-valid) this task establishes whether children sort by perceptual similarities or the formal properties of a shape (e.g., a pre-defined number of straight sides; Satlow & Newcombe, 1998).

Considering the possibility that social interaction is beneficial for children's learning (Kuhl, 2007, 2011), we predicted that participants would learn more when a teacher was present than when they watched a video of the teacher. Moreover, learning complex concepts, like those taught here, may be enhanced by physical action (Agostinho et al., 2015; Dubé & McEwen, 2015; Goldin-Meadow et al., 2009). Thus, we predicted that children in the doing conditions would learn more about shapes' formal properties than those in the observing conditions. Previous research with the sorting task, used here, suggested that children learn more about typical and atypical shapes than non-valid shapes (Fisher et al., 2013). It was therefore possible that the variables manipulated in the training (activity and teacher presence) would interact with shape type in the test session.

Method

Participants

There were 215 children (girls: n=104) aged 57 to 70 months (M=63 months, SD=4 months) recruited via opportunity sampling at five primary schools in England, UK. Four further participants took part in the study but were later excluded (disengaged from the experimental session: n=3; undisclosed learning difficulties: n=1). Children came from both urban and semi-rural middle-class neighbourhoods. More than 50% of children came from well-educated families and 75% were White. Full demographic details of the sample and information about children's screen media use are included in the Supplementary Materials. The University [name] Ethics Committee approved the experiment.

Design

The experiment adopted a mixed factorial design. The between-participant variables were 'activity' (doing, observing) and 'teacher presence' (present, on-screen); the within-participant variables were 'time' (pre-training, post-training) and 'shape-type' (typical, atypical, non-valid). Children were quasi-randomly assigned to one of the four experimental conditions: *doing with teacher-present* (n=56), *observing with teacher present* (n=54), *doing with teacher-on-screen* (n=52), *observing with teacher-on screen* (n=53). The dependent variable was accuracy on the shape-sorting task.

Apparatus and materials

In both teacher-on-screen conditions (doing, observing), the teaching materials were presented on a 13-inch Dell Inspiron convertible 2-in-1laptop set up in a tablet mode (i.e., touchscreen monitor, no visible keyboard). The scripts used in all teaching conditions were based on the materials developed by Fisher et al. (2013). For the example of complete scripts, see Supplementary Materials. We selected three target shapes: rectangle, pentagon and hexagon but excluded triangle, as during piloting of

the sorting task, children's accuracy exceeded 80% for a typical exemplar. Accuracy for the other typical exemplars during piloting ranged from 46% (rectangles and pentagons) to 50% (for hexagons).

Although the content of the *doing* and *observing* scripts was the same, the scripts aimed to differentially engage social and cognitive learning mechanisms (for details see Table 1). Specifically, the *doing* script contained phrases that aimed to elicit children's interest and interaction with either present or on-screen teacher. For example, the teacher said: "Today **I need your help** in discovering the secrets of these shapes. **Let's figure out together** what makes a shape a real shape"(Fisher et al., 2013; Supplemental Materials: Appendix S1; Guided play). In contrast, in the *observing* script, the teacher did not attempt to actively engage a child. Instead, a child was instructed to watch. For example, she said: "Today **I am going** to discover the secret of shapes. **I am going to figure out** what makes a shape a real shape". Moreover, the *doing* script included two hands-on learning tasks for the children: side counting and shape construction. Although these tasks were also included in the observing conditions, rather than doing the task themselves, children watched the teacher do them.

Table 1. Learning mechanisms in each experimental condition (based on the analysis of Strouse et al., 2013).

Learning mechanism	Condition			
	Doing, teacher-present	Observing, teacher-present	Doing, teacher-on-screen	Observing, teacher-on-screen
Engagement & social interaction	*Very high:* uses both gestures (e.g., finger pointing) and language to draw attention and elicit engagement. Social cues created by the presence of the teacher.	*Medium to high:* uses gestures to draw attention. Social cues created by the presence of the teacher.	*Medium:* uses gestures and editing conventions (e.g., zooms and pans) to draw attention; language to draw attention and elicit interaction	*Low to medium:* uses gestures and editing conventions to draw attention. No attempt to elicit social interaction.
Cognitive	*Very high:* repetition, tactile practice, opportunity to make mistakes.	*Medium:* repetition, opportunity to observe faultless practice.	*High:* repetition, touchscreen 2D practice, opportunity to make mistakes.	*Medium:* repetition, opportunity to observe faultless practice.
Feedback	*High:* feedback contingent on the outcome of practice.	*Very low:* no feedback.	*High:* feedback contingent on the outcome of practice.	*Very low:* no feedback.

In the *doing with teacher-on-screen* condition, the teaching materials were created specifically for this study by combining a pre-recorded video footage of a teacher with two on-screen learning tasks using PyCharm as Integrated Development Environment (IDE) in the Python programming platform (version 2.7.13). On-screen learning tasks were developed using Pygame package (version 1.9.3) and the videos in the interactive teaching materials were played using PsychoPy (version 1.84.1). The codes used to program these sessions are available in Supplementary Materials.

To produce a video, a teacher sitting at the table was filmed from a single camera. In the video, she first explained what "makes a shape a real shape" by

following the doing condition script. To explain the properties of a shape, the teacher

used the laminated cards showing the relevant shape arrays and individual shapes.

Each shape array card showed four exemplars: one typical in a *typical* orientation, one

typical in an *atypical* orientation, for example, upside-down and two atypical

exemplars (for an example, see Figure 1a; the full set of shape array cards is included

in the Supplementary Materials). Individual shape cards showed a single typical

exemplar in a typical orientation. The video also showed the teacher counting sides

(Figure 1b & 1c) and making shapes by attaching wooden sticks to a white laminated
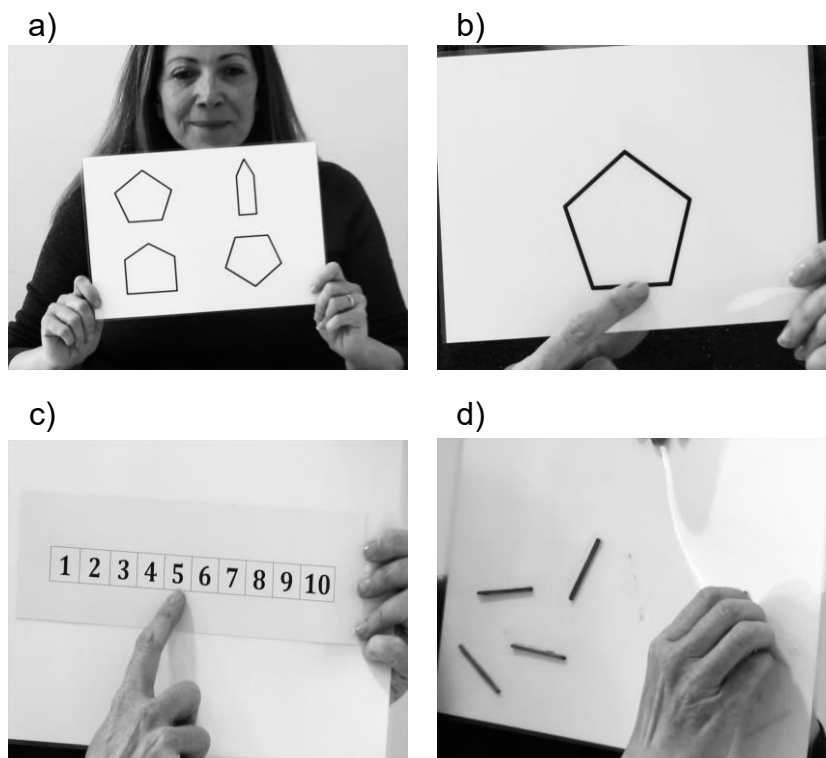
card with a double-sided clear tape (Figure 1d).

Figure 1. The screenshots from the video footage: a) teacher front view, b) teacher

counting sides, c) teacher pointing to the number line, d) teacher constructing a shape.

Subsequently, this footage was combined with two on-screen learning tasks. For example, after the video of the teacher counting the sides of a shape, she asked: "Can you count the sides of a [shape name]? A [shape name] will pop up on your screen. Remember to place your thumb on the first side you count. When you know the number of sides, press the number on your number line." Following these instructions, a typical exemplar of a shape appeared centrally on the screen. Above it was a number line showing digits one to 10 (Figure 2a). When a child had made a response, a balloon showing the pressed digit appeared on the screen (Figure 2b). This was followed by a pre-recorded feedback contingent upon the answer. For example, if a child correctly counted the sides of a rectangle, they would see a video of the teacher saying "You are right. A rectangle has four sides". If a child made no response after 45 s, a video prompt would appear: "Now you count. Remember to place your finger on the first side you count. When you know the number of sides press a number button". If no response were made after this prompt, another video of the teacher counting the sides would appear.

Similarly, the on-screen construction task was introduced immediately after the video of the teacher making a shape from wooden sticks was played. After the teacher had said: "Look I built a [shape]. Now it's your turn. Touch the sticks that will pop up on the screen with your finger and move them around to make a [shape]", randomly arranged virtual sticks appeared on the screen (Figure 2c).
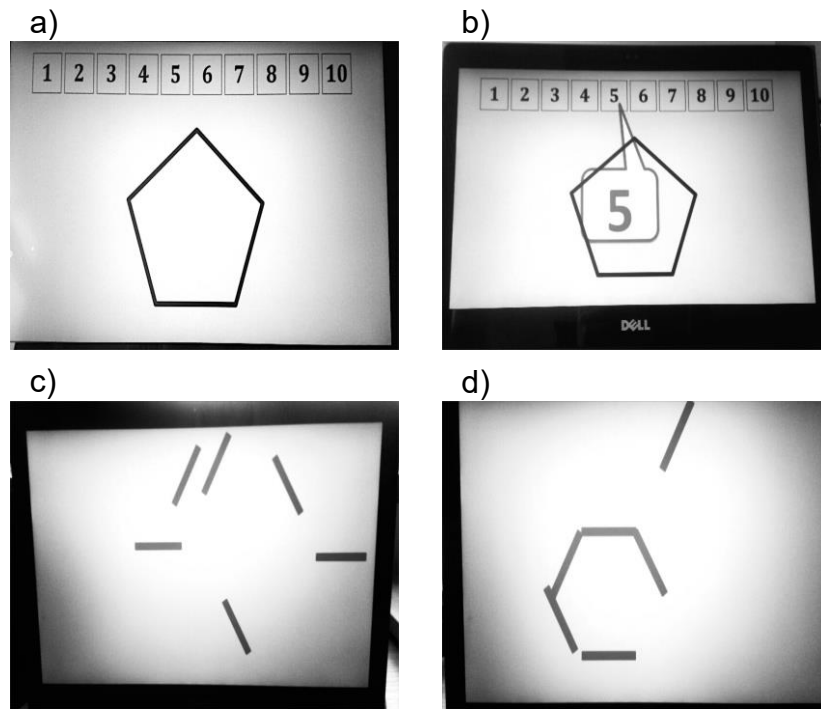
Figure 2. The screenshots showing *doing with teacher-on-screen* condition materials: a) number line before a press, b) number line after a press, c) pre-arranged construction sticks, partially-arranged construction sticks.

The number of sticks matched the number of sides of the target shape. Children manipulated the sticks with their fingers to make a target shape (Figure 2d). Again, a child's response was followed by the contingent feedback. For example, if a child failed to build a correct shape, they would watch a video of the teacher saying: "You've tried hard but that's not quite right. Look at my [shape]" and the teacher showed their laminated card with a shape she had built. If after 30 s a child did not touch any stick on the screen, a video of a prompting message appeared on the screen.

In the *observing with teacher-on-screen* condition, the content of the video was identical, and the teacher used the same teaching aids. However, she followed the observing script.

The teacher-present teaching materials had the same content and used the same teaching aids as both teacher-on-screen conditions. The teaching in the *observing with teacher-present* condition consisted of the experimenter delivering a teaching session based on the observing script. The doing condition mirrored the screen equivalent and combined the live teaching with two hands-on learning tasks. In the counting task, children were instructed to count the sides of a shape shown on an A5 laminated card and then to "press" the correct number on the number line presented on an A4 laminated card. In the construction task, the children used wooden sticks to build a shape.

In the *doing with teacher-present* and *doing with teacher-on-screen* conditions, the children were exposed to each activity twice: when they watched the experimenter count or build and when they did the task themselves. To match the amount of exposure to each shape in the observing conditions, the experimenter completed the counting and construction tasks twice.

For the sorting task, the modified examples of rectangles, pentagons and hexagons used in previous research (Fisher et al., 2013; Satlow & Newcombe, 1998) were presented individually on A5 laminated cards. There were 10 exemplars of each shape: three typical, three atypical and four non-valid. All typical exemplars were convex shapes with equal sides (except the rectangle) and equal interior angles but varied in size and orientation. The atypical examples were also convex shapes but had unequal sides and interior angles. The non-valid shapes included two shapes with one more side than the target shape (one concave and one convex), one shape with fewer sides than the target shape and one shape with a curved side (for details see Supplementary Materials).

Procedure

The experiment took place in a quiet area that was separate from the main classroom area. Children were tested individually. First, they completed a short warm-up task followed by the shape-sorting task. In the shape-sorting task, the experimenter placed two cardboard boxes in front of a child. One box was labelled with a green tick and the other with a red cross. The experimenter explained that the task was to put all the "real" shapes in the box with a green tick, and all the "fake" shapes in the box with a red cross. Thirty shape cards were facing down on the table and, in each trial, the experimenter drew one card, showed it to the child and asked: "Is this is a real [shape name] or a fake [shape name]?" The order of shape presentation was the same for all children (rectangles-pentagons-hexagons), but the cards showing different shape types (typical, atypical, non-valid) were arranged randomly.

When children completed the sorting task, in the teacher-present conditions, the experimenter proceeded with the relevant training session. In the teacher-on-screen conditions, the experimenter introduced the video or a tablet "game" and instructed children to watch and listen carefully to the teacher on the screen. No further instructions were given when children were watching the video or playing the game. After the training session, each child completed the second shape-sorting task. Upon completing the session, each child received a small reward for taking part.

Results

The pre- and post-training accuracy for the shape-sorting task in each condition is shown in Table 2. At pre-test, there was no difference in task performance between the experimental groups (all p-values>.05).

First, the pre- and post-training scores were analysed in a mixed analysis of variance (ANOVA) with activity (doing, observing) and teacher presence (present, on-screen) as the between-subject variables and time (pre- and post-training) as the within-subject variable. Second, we calculated learning as a difference between pre- and post-training accuracy and conducted a further mixed ANOVA on learning scores with shape type (typical, atypical, non-valid) as the within-subject variable. When the assumption of sphericity was not met, we adjusted the degrees of freedom using the Greenhouse-Geisser correction. Finally, when performing multiple comparisons following a significant interaction, we controlled the family wise error rate using a bootstrap-based resampling method. We drew 10,000 bootstrap samples to produce bias-corrected and accelerated 95% confidence intervals (BCa 95% CIs).

Table 2. Mean (SD) percentage accuracy on the shape-sorting task by shape type pre- and post-training in each condition.

Shape type	Condition			
	Doing with teacher-present	Observing with teacher-present	Doing with teacher-on-screen	Observing with teacher-on-screen
Pre-training				
Typical	75.4% (18.4)	71.8% (25.1)	76.5% (22.1)	75.3% (24.3)
Atypical	16.5% (22.0)	14.4% (16.8)	22.0% (23.4)	21.0% (22.2)
Non-valid	87.2% (18.7)	87.3% (14.0)	81.1% (23.5)	85.8% (14.8)
Post-training				
Typical	84.1% (18.7)	78.8% (27.7)	84.0% (19.6)	79.7% (22.8)
Atypical	53.2% (36.4)	48.3% (36.1)	47.0% (38.6)	44.0% (32.9)
Non-valid	83.2% (21.2)	79.5% (23.4)	74.2% (18.9)	79.5% (20.4)

The first mixed ANOVA showed a significant main effect of Time, $F(1, 211)=111.36$, $p<001$, $\eta_p^2=.345$. This was qualified by a significant Time x Teacher's Presence interaction, $F(1, 211) = 4.90$, $p=.028$, $\eta_p^2=.023$. However, there were no significant effects of Teacher's Presence (.676), Activity ($p=.345$), nor significant

Time x Activity interaction (p=.326). These results suggest that although training

improved all children's shape knowledge, the group that learned with a teacher

present benefited the most.

The second mixed ANOVA did not produce a significant main effect of

Activity (p=.636). There was, however, a main effect of Shape Type,

$F(1.83,385.32)=70.24$, p<.001, $\eta_p^2=.250$, which was qualified by a significant Shape

Type x Teacher presence interaction, $F(1.83,385.32)=3.40$, p=.038, $\eta_p^2=.016$ (Figure

3). Follow-up, bootstrapped, independent samples t-tests showed that for atypical

shapes, children learned significantly better with the teacher present than from a

video, $t(213)=2.32$, p=.022, BCa 95%CIs: 1.60 to 21.04%. For typical and non-valid

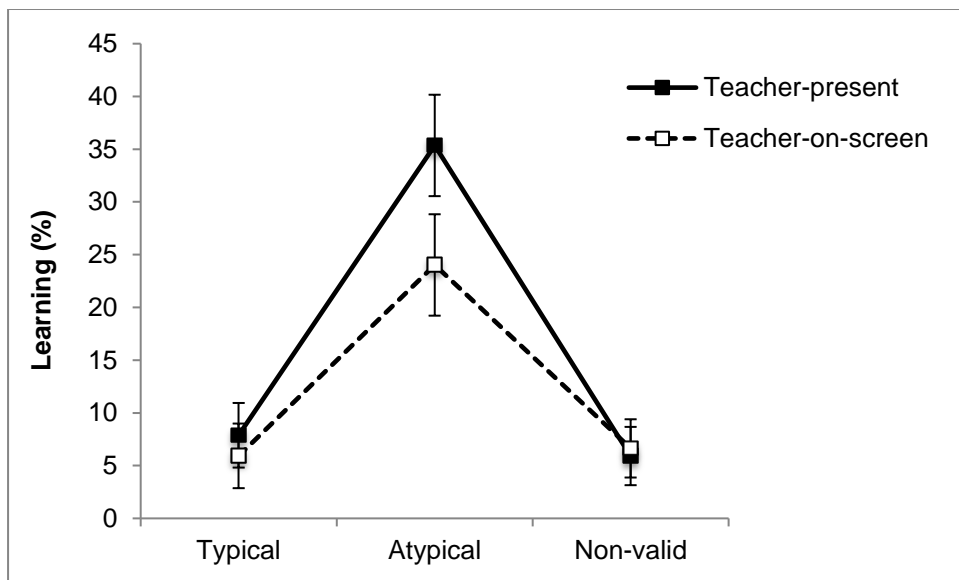shapes, teacher presence made no difference to learning (p=.536 and p=.921,

respectively).

Figure 3. Learning for each shape type either with the teacher present or from a

video. Errors bars represent 95% CIs.

17

Discussion

This study investigated what makes young children learn. Specifically, we manipulated a teacher's presence (present vs. on-screen) and activity (doing vs. observing) to compare children's learning of a difficult mathematical concept (the formal definitions of shapes). Our first hypothesis predicted that children would learn more when a teacher was present than they did when watching a video of one (e.g., Kuhl, 2007, 2011). The data supported this prediction. Although children did learn in all conditions, they learned most when the teacher was present. This effect replicates findings from language research (e.g., DeLoache et al., 2010; Krcmar, 2011; Krcmar et al., 2007; Kuhl et al., 2003; Roseberry et al., 2009), in which infants and children were found to learn more when a teacher was present. Thus, our data support the *social learning hypothesis* (Kuhl, 2007, 2011), which proposes that social interaction is fundamental to language learning. However, more detailed analysis showed that teacher's presence only mattered when children learned about atypical shapes. This may indicate that while screen learning improves children's understanding of simple content, to enhance learning concepts that are difficult relative to cognitive maturity a teacher's presence is more important.

Applying Kuhl's (2007, 2011) theory, the presence of a teacher might have improved our children's learning in three ways. First, the social engagement provided by the teacher might have increased children's overall attention and arousal during the training session to enhance learning. Second, children might have used the teacher's non-verbal behaviour to obtain referential information. Specifically, the teacher might have used cues such as gaze direction, hand gestures or whole body movements (Sebanz, Bekkering, & Knoblich, 2006) to emphasise the key definitional features of atypical shapes. In turn, this might have fostered learning by enhancing children's

attention to specific relevant content (Kuhl, 2007). Such referential non-verbal cues were also present in the videos. However, video training relies less on social cues and more on editing techniques to draw a learner's attention to the relevant content (Hirsh-Pasek et al., 2015). Thus, extracting referential details from the on-screen teacher could be more difficult (Kuhl, 2007). While 5-year-olds may understand the symbolic meaning of editing, they may not be able to access the relevant content as effortlessly as when interacting with a real teacher.

Finally, through millions of years of evolution, brain mechanisms underpinning gestural and verbal communication that support learning have evolved to extract meaningful information from the behaviour of a *live* human (Dickerson, Gerhardstein, & Moser, 2017; Kuhl, 2011). Arguably, the putative human 'mirror neuron system' may allow one to link their own behaviour with actions (Rizzolatti & Craighero, 2004) and speech (Iacoboni & Dapretto, 2006) of their social partner. In this study, the *teacher-present* conditions provided a natural platform for face-to-face live interactions, thus, allowing the potential activation of the mirror neuron system. Conversely, in the *doing with teacher-on-screen* condition, these social interactions were mediated by the screen. It is therefore possible that the presence of spatial and temporal discontinuity in child-teacher interactions that may be present whilst communicating via digital media (Dickerson et al., 2017), disrupted the timing of neural activation and, consequently, reduced learning.

Turning to our second hypothesis, we predicted that children in the doing conditions would learn more than those in the observing conditions (Agostinho et al., 2015; Dubé & McEwen, 2015; Fisher et al., 2013; Goldin-Meadow et al., 2009). In this study, the doing conditions involved hands-on practice with either physical shapes or their touchscreen equivalents (versus watching in the observing conditions).

Children were also given feedback on the accuracy of their performance (with no feedback in the observing conditions). However, the addition of these features to the training session did not improve children's learning, as measured on a subsequent shape-sorting task. This contrasts with some previous findings, in which, joint exploration, feedback and interaction with either physical or touchscreens objects (Agostinho et al., 2015; Dubé & McEwen, 2015; Fisher et al., 2013) was associated with more effective learning (although see Aladé et al., 2016; Schroeder & Kirkorian, 2016).

One explanation for the absence of an effect of activity, in the current study, is the way we manipulated our training. Our aim was to ensure that the four training conditions only differed in activity and teacher presence, and that participants within each condition received identical training. Whilst this achieved effective experimental control, it may have limited children's opportunity for *exploration* during training. For example, when building shapes in the doing conditions, children could only use a predetermined number of equilateral construction sticks. This, in turn, might have limited children's opportunities to discover non-defining shape features (i.e., size, orientation, skewness; Aslan & Arnas, 2007; Satlow & Newcombe, 1998), and ultimately reduced learning (Antle, 2009). Similarly, the nature of feedback was precisely defined by the script and did not allow for the natural back-and-forth often associated with this process. Future research should investigate which variables (e.g., physical exploration or flexible social interaction) mediate the effects of hands-on experience and feedback on learning.

Another explanation might be that, because all children were tested with a physical task, those who trained with a tablet were tested in different context to that used in training. This meant that children in the *teacher-on-screen* condition changed

from an on-screen to a physical task, whereas children in the *teacher-present* condition remained with the physical task. However, as noted in the introduction, the use of 2D stimuli in all conditions reduced the effect of this change. More generally, the issue of transferring learning from an on-screen to a physical task is important and should be addressed in future research. This is because many schools use tablets as learning aids, but test children with traditional pen-and-paper methods.

We must acknowledge that while producing own learning materials allowed us to carefully manipulate children's learning experience, it may have reduced the generalizability of the findings. The majority of commercially available educational apps do not contain a video of a teacher delivering training. Instead, apps gamify learning experience and prompt children to perform an action with simple text or pictorial instructions (e.g., tap or drag here). However, apps using interactive video tutorials are not entirely absent from the digital education market (e.g., Khan Academy Kids– a free online learning platform). Future research should identify which type of apps (game-like or interactive video tutorials) maximise young children's learning.

In conclusion, our findings emphasise the importance of a teacher to children's learning of difficult concepts. They are consistent with Kuhl's (2007) social learning hypothesis, which identifies specific links between social input and language learning (Kuhl, 2011), while recognising that its value may not be restricted to this domain. Despite current enthusiasm for the potential learning opportunities created by touchscreen devices (e.g., Huber et al., 2016; Wang, Xie, Wang, Hao, & An, 2016), we found no support for the proposal that screen interactivity fosters learning beyond what could be achieved by passive viewing. Considering the growing popularity of using interactive devices as learning aids both at home and in the classroom, further

research is essential to establish under which conditions they may aid children's

learning.

References

Agostinho, S., Tindall-Ford, S., Ginns, P., Howard, S. J., Leahy, W., & Paas, F. (2015). Giving learning a helping hand: Finger tracing of temperature graphs on an iPad. *Educational Psychology Review, 27*, 427-443.

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.

Alibali, M. W., & Nathan, M. J. (2012). Embodiment in mathematics teaching and learning: Evidence from learners' and teachers' gestures. *Journal of the Learning Sciences, 21*, 247-286.

Antle, A. N. (2009). Lifelong interactions. Embodied child computer interaction: Why embodiment matters. *interactions, 16*, 27-30.

Aslan, D., & Arnas, Y. A. (2007). Three‑to six‑year‑old children's recognition of geometric shapes. *International Journal of Early Years Education, 15*, 83-104.

Barr, R. (2010). Transfer of learning between 2D and 3D sources during infancy: Informing theory and practice. *Developmental Review, 30*(2), 128-154.

Baydar, N., Kağitçibaşi, Ç., Küntay, A. C., & Gökşen, F. (2008). Effects of an educational television program on preschoolers: Variability in benefits. *Journal of Applied Developmental Psychology, 29*, 349-360.

DeLoache, J. S., Chiong, C., Sherman, K., Islam, N., Vanderborght, M., Troseth, G. L., . . . O'Doherty, K. (2010). Do babies learn from baby media? *Psychological science, 21*, 1570-1574.

Dickerson, K., Gerhardstein, P., & Moser, A. (2017). The Role of the Human Mirror Neuron System in Supporting Communication in a Digital World. *Frontiers in Psychology, 8*, 698.

Dubé, A. K., & McEwen, R. N. (2015). Do gestures matter? The implications of using touchscreen devices in mathematics instruction. *Learning and Instruction, 40*, 89-98.

Fisher, K. R., Hirsh-Pasek, K., Newcombe, N., & Golinkoff, R. M. (2013). Taking shape: Supporting preschoolers' acquisition of geometric knowledge through guided play. *Child Development, 84*, 1872-1878.

Galetzka, C. (2017). Commentary: Mobile and interactive media use by young children: The good, the bad, and the unknown. *Frontiers in Psychology, 8:461*.

Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. A. (2009). Gesturing gives children new ideas about math. *Psychological science, 20*, 267-272.

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*, 3-34.

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.

Iacoboni, M., & Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nature Reviews Neuroscience, 7*, 942-951.

Kontra, C., Goldin-Meadow, S., & Beilock, S. L. (2012). Embodied learning across the life span. *Topics in Cognitive Science, 4*, 731-739.

Kostyrka-Allchorne, K., Cooper, N. R., & Simpson, A. (2017). The relationship between television exposure and children's cognition and behaviour: A systematic review. *Developmental Review, 44*, 19-58.

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

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

Kuhl, P. K. (2007). Is speech learning 'gated'by the social brain? *Developmental science, 10*, 110-120.

Kuhl, P. K. (2011). Social mechanisms in early language acquisition: Understanding integrated brain systems supporting language. In J. Decety & J. T. Cacioppo (Eds.), *The Oxford handbook of social neuroscience* (pp. 649-667). New York: Oxford University Press, Inc. .

Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences, 100*, 9096-9101.

Linebarger, D. L., Kosanic, A. Z., Greenwood, C. R., & Doku, N. S. (2004). Effects of viewing the television program between the lions on the emergent literacy skills of young children. *Journal of educational psychology, 96*, 297-308.

Lovato, S. B., & Waxman, S. R. (2016). Young children learning from touch screens: taking a wider view. *Frontiers in Psychology, 7:1078*.

Miklashevsky, A. A., & Fischer, M. H. (2017). Commentary: Down with Retirement: Implications of Embodied Cognition for Healthy Aging. *Frontiers in Psychology, 8:599*.

Moyer, P. S. (2001). Are we having fun yet? How teachers use manipulatives to teach mathematics. *Educational Studies in mathematics, 47*, 175-197.

Nield, D. (2016). 15 memorable milestones in tablet history. Retrieved from http://www.techradar.com/news/mobile-computing/10-memorable-milestones-in-tablet-history-924916

Ofcom. (2017). *Children and parents: Media use and attitudes report*. Retrieved from

Radesky, J. S., Schumacher, J., & Zuckerman, B. (2015). Mobile and interactive media use by young children: The good, the bad, and the unknown. *Pediatrics, 135*, 1-3.

Radich, J. (2013). Technology and interactive media as tools in early childhood programs serving children from birth through age 8. *Every Child, 19*, 18-19.

Reiß, M., Krüger, M., & Krist, H. (2019). Theory of mind and the video deficit effect: Video presentation impairs children's encoding and understanding of false belief. *Media Psychology, 22*(1), 23-38.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual review of neuroscience, 27*, 169-192.

Roseberry, S., Hirsh-Pasek, K., Parish-Morris, J., & Golinkoff, R. M. (2009). Live action: Can young children learn verbs from video? *Child Development, 80*, 1360-1375.

Satlow, E., & Newcombe, N. (1998). When is a triangle not a triangle? Young children's developing concepts of geometric shape. *Cognitive development, 13*, 547-559.

Schroeder, E. L., & Kirkorian, H. L. (2016). When seeing is better than doing: Preschoolers' transfer of STEM skills using touchscreen games. *Frontiers in Psychology, 7:1377*.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in cognitive sciences, 10*, 70-76.

Spitzer, M. (2013). To swipe or not to swipe?—The question in present-day education. *Trends in Neuroscience and Education, 3*, 95-99.

Strouse, G. A., O'Doherty, K., & Troseth, G. L. (2013). Effective coviewing: Preschoolers' learning from video after a dialogic questioning intervention. *Developmental Psychology, 49*, 2368.

Wang, F., Xie, H., Wang, Y., Hao, Y., & An, J. (2016). Using touchscreen tablets to help young children learn to tell time. *Frontiers in Psychology, 7:1800*.

Wright, J. C., Huston, A. C., Murphy, K. C., St Peters, M., Pinon, M., Scantlin, R., & Kotler, J. (2001). The relations of early television viewing to school readiness and vocabulary of children from low‐income families: The early window project. *Child Development, 72*, 1347-1366.