EDC Education Development Center

SRI Education™
A DIVISION OF SRI INTERNATIONAL

Ready To Learn **Research** November 2019

Early Science & Engineering

The Impact of *The Cat in the Hat Knows a Lot About That!* on Learning

About EDC

Education Development Center (EDC) is a global nonprofit that advances lasting solutions to improve education, promote health, and expand economic opportunity. Since 1958, we have been a leader in designing, implementing, and evaluating powerful and innovative programs in more than 80 countries around the world.

About SRI

SRI Education, a division of SRI International headquartered in Menlo Park, California, is tackling the most complex issues in education and learning to help students succeed. We work with federal and state agencies, school districts, major foundations, nonprofit organizations, and international and commercial clients to address risk factors that impede learning, assess learning gains, and use technology for educational innovation.

Authors

Todd Grindal, Megan Silander, Sarah Gerard, Tiffany Maxon, Elisa Garcia, Naomi Hupert, Phil Vahey, Shelley Pasnik

Contributing Researchers

Alexandra Adair, Laura Bartel, Wei-Bing Chen, Kirby Chow, Claire Christensen, Kate Ferguson, Annie Fikes, Stephany Garcia, Shari Gardner, Pamela Genao Reyes, Sara Gracely, Jaime Gutierrez, Saramax Guttman, Cindy Hoisington, Joy Kennedy, Sarah Kim, Janna Kook, Elise Levin-Guracar, Jennifer Nakamura, Lucy Nelson, Jenna Nguyen, Min-Kyung Park, Samantha Peyton, Lan Pham, Alexia Raynal, Daniela Saucedo, Kate Sinclair, Regan Vidiksis, Laura Zimmermann

Acknowledgements

We are extremely grateful to the participating families who took the time from their busy schedules to participate in the study. Without their participation, this study would not have been possible. We also would like to thank Daryl Greenfield and his research team at the University of Miami for allowing the use of the Lens on Science assessment and for their support throughout the study. Finally, we also wish to thank PBS advisors Kim Brenneman, Karen Worth, and Sara Sweetman for their feedback on the study.

Suggested citation:

Grindal, T., Silander, M., Gerard, S., Maxon, T., Garcia, E., Hupert, N., Vahey, P., Pasnik, S. (2019). *Early Science and Engineering: The Impact of* The Cat in the Hat Knows a Lot About That! *on Learning*. New York, NY, & Menlo Park, CA: Education Development Center, Inc., & SRI International.

About the Ready To Learn Initiative

The Ready To Learn Initiative is a cooperative agreement funded and managed by the U.S. Department of Education's Office of Elementary and Secondary Education. It supports the development of innovative educational television and digital media targeted to preschool and early elementary school children and their families. Its general goal is to promote early learning and school readiness, with a particular interest in reaching children from low-income households. In addition to creating television and other media products, the program supports activities intended to promote national distribution of the programming, effective educational uses of the programming, community-based outreach, and research on educational effectiveness.

The contents of this research report were developed under a grant from the Department of Education. However, those contents do not necessarily represent the policy of the Department of Education, and you should not assume endorsement by the Federal Government. [PR/Award No. U295A150003, CFDA No. 84.295A]

Design: EDC Digital Design Group
Photography: Burt Granofsky, EDC

Contents

Executive Summary

Objectives

Digital media—such as television, videos, apps, and digital games—represent a potentially powerful tool to support children's science and engineering learning at scale. In this study, we test whether providing young children and their families with science and engineering-focused digital media leads to improvements in children's science- and engineering knowledge and practices.

Resources Tested

This report examines *The Cat in the Hat Knows a Lot About That!*™, a PBS KIDS multi-platform media property based on "The Cat in the Hat's Learning Library" book series by Random House and Dr. Seuss Enterprises. The television series is produced by Portfolio Entertainment, and the digital games and offline activities are produced by Random House. The Public Broadcasting Service (PBS), in partnership with the Corporation for Public Broadcasting (CPB), developed the media resources researched in this report as part of the 2015–2020 Ready To Learn Initiative, funded through the U.S. Department of Education. *The Cat in the Hat Knows a Lot About That!* resources tested were 28 digital videos, five digital games, and hands-on activities for parents and children.

28 **videos** (3 hours and 43 minutes of content) in a video app

5 **digital games** in *The Cat in the Hat Builds That* app

20 **game-aligned guides** for hands-on activities embedded in the game app

3 **printed descriptions** of science and engineering activities from the website

Methods

The study team randomly assigned 454 children from low-income families to receive either *The Cat in the Hat Knows a Lot About That!* resources or an alternative treatment. The team collected information on children's science and engineering knowledge and practices prior to randomization and then again eight weeks later.

Who participated in the study?

A diverse sample of families, all living in low-income households.

454

Children:
Ages 4 and 5 years

Key Findings

Exposure to *The Cat in the Hat Knows a Lot About That!* resources had meaningful impacts on 4- to 5-year-old children's physical science knowledge and their ability to engage with science and engineering practices.

Exposure to *The Cat in the Hat Knows a Lot About That!* media resources had a clear positive impact on children's

1. understanding of physical science concepts related to the two physical science core ideas of matter and forces:[1] **(a)** the role of material properties (strength and length) and forces in structural stability and **(b)** the role of material properties (texture) and forces (friction) on movement down an incline; and
2. interest and engagement in science.

We also found suggestive evidence that exposure to *The Cat in the Hat Knows a Lot About That!* had a positive impact on **(a)** children's wider understanding of physical science concepts and science and engineering practices; **(b)** their ability to sort objects by size, color, shape, and use; and **(c)** their use of science-related vocabulary. These impacts were small and approached—but did not reach—the conventional threshold for statistical significance.

Implications

The results of this study indicate that educational media designed to focus on critical science and engineering concepts and skills can help young children understand science and engineering concepts and practices. Second, large impacts on two of the three performance-based assessments indicate that children's experiences manipulating materials in a digital context can transfer to understanding of those practices and concepts in the physical world. This finding provides a substantial contribution to overall research literature on how children learn from digital media. These findings are especially important given the relative scalability and low per-child cost of media-based interventions in the context of other typically resource-heavy science interventions.

[1] These concepts are related to two core ideas derived from the PBS KIDS Science Framework, which is aligned with the Next Generation Science Standards: (1) matter and Its interactions and (2) motion and stability: forces and interactions.


PERSIST

Introduction

Young children are **naturally curious**. They delight in asking questions about the world and how things work, generating explanations, using models, and making and revising predictions.

(Fusaro & Smith, 2018; Gerde, Schachter, & Wasik, 2013; National Research Council, 2012)

These behaviors are the first steps in developing important science and engineering knowledge and practices as well as in developing critical thinking, language, mathematical reasoning, executive function, and persistence (Bustamante, White, & Greenfield, 2017; French, 2004; Kuhn, 2011; Nayfeld, Fuccillo, & Greenfield, 2013; Peterson & French, 2008; Wright & Neumann, 2014). Children who engage in science are more likely to be interested in science as adolescents and are more likely to participate in science-related careers later in life (Alexander, Johnson, & Kelley, 2012; Patrick, Mantzicopoulos, & Samarapungavan, 2009).

However, many young children have limited opportunities to explore science and engineering at home or in preschool settings. Four- and 5-year-old children in preschool spend far less time learning science content than they do other subjects, such as language and literacy (Greenfield et al., 2009; Piasta, Pelatti, & Miller, 2014). Many preschool educators are not trained to support children's science and engineering explorations (Greenfield et al., 2009). And at home, parents[2] often struggle to answer their child's questions about the world around them, and many lack the confidence and resources to seize everyday opportunities that could deepen their child's learning and curiosity about how things work (Silander et al., 2018).

Digital media—television, video, apps, digital games, and so on—represent a potentially powerful tool to support children's science learning (Mares & Pan, 2013; Schroeder & Kirkorian, 2016). Media can model ways of talking about and doing science and can engage families who might otherwise perceive science as too complex or not interesting (Lu, Buday, Thompson, & Baranowski, 2016; Travis, 2017; Troseth, Saylor, & Archer, 2006; Young et al., 2012). Digital media can allow

[2] We use the term "parent" in a broad sense, as our sample includes guardians as well as parents.

Prior research suggests a need to provide young children and their families with more high-quality science- and engineering-focused media resources.

children to explore and manipulate scientific phenomena and simulations that might be time consuming, difficult, or unsafe to explore in the physical world (Rutten, Van Joolingen, & Van der Veen, 2012). Media is most likely to support learning and development if it features relatable characters and stories (Bandura, 1965; Lauricella, Gola, & Calvert, 2011; Linebarger, Brey, Fenstermacher, & Barr, 2017), is cognitively activating and engaging, supports meaningful and socially interactive learning experiences, and is guided by specific learning goals (Hirsch-Pasek et al., 2015). However, even with well-designed media, transferring knowledge from videos and games to the real world[3] is challenging, particularly for the youngest media users (Bonus & Mares, 2019; Troseth, 2010). Related to this challenge, media's potential to educate is strengthened when children and parents watch and play together, and joint media engagement can be a particularly powerful tool for learning (Rasmussen et al., 2016; Strouse, O'Doherty, & Troseth, 2013).

Educational media's ability to deliver consistent messaging to a wide and varied audience across large geographical areas makes educational media interventions highly scalable at an extremely low cost relative to other person-to-person early childhood interventions (Kearney & Levine, 2019). Children under the age of 8 years spend an average of two hours a day using or watching onscreen media, including television, movies, and apps (Rideout, 2017). Although the potential exists for media to change perceptions and influence outcomes related to science and engineering, a recent national survey of parents found that just over half of parents were satisfied with the science learning media resources available to their children, and even fewer reported that their child "learned a lot" from science media (Silander et al., 2018). These findings suggest a need to provide young children and their families with more high-quality science- and engineering-focused media resources.

To meet this need, the Public Broadcasting Service (PBS), in partnership with the Corporation for Public Broadcasting (CPB), developed the

[3] We use the term "real world" to refer to activities that are not mediated by the tablet computer.

The Cat in the Hat Knows a Lot About That! *(Cat in the Hat)* media resources for young children and their parents and caregivers through the 2015–2020 Ready To Learn Initiative, which is funded through the U.S. Department of Education. The Ready To Learn Initiative brings free educational television and digital media resources to children ages 2–8, promoting early learning and school readiness, with an emphasis on supporting children from low-income, underserved communities. Developing a deeper understanding of how national media and the network of local public media stations can support children and family learning at scale drives the CPB-PBS Ready To Learn Initiative. As the research and evaluation team for Ready To Learn, Education Development Center (EDC) and SRI Education (SRI) measure the impact of PBS KIDS digital and hands-on resources on young children's learning.

No large-scale rigorous experimental studies have examined the effectiveness of the existing science- and engineering-focused educational media on children's learning. This study examines how exposure to the *Cat in the Hat* videos, digital games, and hands-on activities affects children's understanding of science and engineering concepts and practices.

The Cat in the Hat Knows a Lot About That! Resources

The Cat in the Hat Knows a Lot About That! is a PBS KIDS multi-platform media property based on "The Cat in the Hat's Learning Library" book series by Random House and Dr. Seuss Enterprises. This study focused specifically on resources developed for the third season of *The Cat in the Hat Knows a Lot About That!*, which PBS developed in partnership with CPB under the 2015–2020 Ready To Learn Initiative. The content for this season is aligned with the K–12 Next Generation Science Standards (NGSS) (National Research Council, 2013) and the Head Start Early Learning Outcomes Framework (U.S. Department of Health and Human Services, 2015). This third season of the television series, produced by Portfolio Entertainment, along with the digital games, apps, interstitials, and real-world activities produced by Random House, support science and engineering learning by modeling science and engineering practices and language and exploring science and engineering content through narrative stories and interactive experiences. The third season specifically focuses on physical science and engineering content, practices, and vocabulary, which are topics that work well for visual storytelling (Sweetman, Mirkin, Lund, & Bishop, 2018). The *Cat in the Hat* content developers selected one or two focal concepts to integrate with science and engineering practices in each episode or game.

At the beginning of each *Cat in the Hat* video adventure, two young characters, Sally and Nick, pose a question or define a problem about the natural and human-designed world they have encountered in their own backyard. The characters then join the lead character, *Cat in the Hat*, on a science-related adventure, such as visiting *Spansylvania* to build bridges, going to Gongolia to find the right size gong to make different sounds, or going to the Frictionarium to investigate why Sally has lost her "slidey-ness" and is stuck on the slide. Guided by the Cat, the children engage in science inquiry and engineering design. They ask questions and define problems, make observations, make predictions, plan and conduct investigations, collect data, make discoveries, design solutions to problems, and generate and discuss ideas about how the world works.[4]

In addition to the videos, the content developers created five digital games and guidance for five sets of real-world activities within the PBS *The Cat in the Hat Builds That* app. The digital games are designed to support science and engineering learning by providing opportunities for children to explore and manipulate scientific phenomena and simulations. The real-world activities are each aligned with one of the five games and provide an opportunity for children (ideally with older family members) to engage in related real-world science and engineering activities. Printable PDFs with guidance for additional real-world activities also are available on the *Cat in the Hat* website. Additional details about the *Cat in the Hat* resources, including individual media assets, are listed below in the Study Materials section.

[4] Partially adapted from PBS KIDS, https://www.pbs.org/parents/shows/cat-in-the-hat/about/


rebend all crease (score)
ld - Remove at perforati
r d'avance toutes les ma
uvercle - Enlever aux pe

Study Design and Methods

Research Questions

The study team sought to answer four primary research questions:[5]

What is the impact of providing eight weeks of access to the *Cat in the Hat* resources on low-income 4- to 5-year-old children's understanding of

1. physical science concepts and science and engineering practices;
2. the role of material properties (strength and length) and forces in structural stability;
3. the role of material properties (texture) and forces (friction) on movement down an incline: and
4. how objects can be sorted based on their material properties and uses.

We also examined the following exploratory questions:

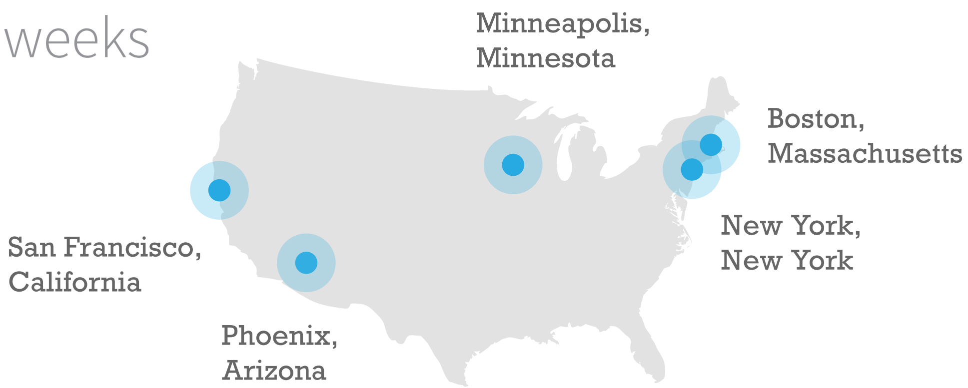
5. In what ways do 4- to 5-year-old children and their parents engage with (that is, discuss, practice, try out in the world, show siblings, and so on) the *Cat in the Hat* resources?
6. In what ways, if at all, is usage of the *Cat in the Hat* resources associated with 4- to 5-year-old children's physical science and engineering knowledge?
7. In what ways, if at all, does the impact of the *Cat in the Hat* resources on child outcomes differ for relevant subgroups of children, including gender, ethnicity, children from dual language homes, parent education level, high-achieving children (as measured by the pre-test), and children in formal preschool settings?

[5] The study team published the study design and analysis plans in the Registry of Efficacy and Effectiveness Studies (REES) prior to beginning data collection. Preregistration of study analysis is done to reduce the possibility of study team members manipulating the analyses to produce positive results. The design and analysis practices described in this report are consistent with the information in the preregistry document. Note that the wording of research questions was changed slightly to improve accessibility to a broader audience. REES 1627.1v2; https://sreereg.icpsr.umich.edu/framework/pdf/index.php?id=2294

Recruitment

The study took place over an eight-week period during winter and spring 2019.

8 weeks



The study team recruited families in five locations across the United States: Boston, Massachusetts; Minneapolis, Minnesota; New York, New York; Phoenix, Arizona; and San Francisco, California. Working in collaboration with Drive Research, a qualitative research and recruiting firm, the study team contacted parents through existing email lists and advertisements on social media. Drive Research contacted interested families by phone to provide additional information about the study and to confirm that potential participants met the following criteria:

- Child was born between September 2013 and October 2014 (age 4 years and 3 months to 5 years and 6 months at baseline).
- Child is fluent in English.
- Family is low income.[6]
- At least one parent is proficient in English or Spanish.
- Parent indicated that child could participate in game-like activities for up to 35 minutes (the duration of the pre-test).

Recruiters made these screening phone calls in English and in Spanish. Families that met these criteria and remained interested were scheduled to participate in two meetings with the study team. During the first meeting, members of the study team informed parents that the purpose of the study was to

[6] For this study, a family qualified as low income if they had an annual household income at or below the U.S. Department of Housing and Urban Development threshold to receive a Section 8 housing voucher for a family of four. This was $81,100 in Boston, $71,900 in Minneapolis, $83,450 in New York, $55,300 in Phoenix, and $117,400 in San Francisco.

understand how young children learn from digital media, that they would receive a tablet computer and $125 in incentives ($25 at the first meeting and $100 at the final meeting), and that they would be asked to complete parent surveys and permit their children to participate in assessments of their early learning knowledge and skills. A study team member then obtained the parents' consent to participate in the study.[7] Table 1, below, displays the number of participants who successfully completed the first meeting at each site. The second meeting occurred eight weeks later and included post-test data collection. These procedures are described in detail in Appendix A.

Table 1. Study Participants by Site

	N	% of total sample
Boston	29	6.4%
Minneapolis	121	26.7%
New York	106	23.3%
Phoenix	116	25.6%
San Francisco	82	18.1%

[7] At the time of consent, parents also were given the option to participate in a second study. Activities associated with the second study took place after the completion of this study. Families could participate in the first study whether or not they chose to participate in the second. The second study will be described in a separate report.

Random Assignment

The study team randomly assigned participants to the treatment or control group by provisioning all study tablet computers in a random sequence with control or treatment materials before the first data collection visit. Once parents completed the study enrollment process, they received either a treatment- or control-provisioned tablet. A member of the study team distributed tablets sequentially, and the study team members who determined the order of the participants were not aware of the tablet sequence. This process ensured that assignment to condition was entirely random and could not be influenced by the participants or the study team. To mitigate possible biases in data collection and analyses, the study procedures throughout the study ensured that assessors and data analysts were not aware of the study participants' assignment to condition.

Sample

The children who participated in the study were an average of 4 years, 11 months old. The sample included slightly more girls (54%) than boys. Approximately one third (35%) of participating children were identified by their parent as White, 19% as Hispanic, 18% as Black or African American, 3% as Asian, and 23% as multiracial. All families met the local criteria for qualifying as low-income, and approximately half (51%) lived in households with annual incomes of less than $50,000. Three quarters (75%) of the children spoke English as a primary language, and 10% of parents indicated that their child received disability services (via an individualized education program or a 504 plan). Table 2, below, displays the demographic characteristics of the families in the overall sample and by condition. We did not observe any statistically significant ($p < .10$) differences between treatment- and control-assigned children or families on any of these demographic characteristics, and in no case did differences exceed the What Works Clearinghouse threshold (ES = 0.25) for equivalent samples.

Table 2. Child Science Knowledge and Demographic Characteristics, Overall and by Condition

	Total Sample (N = 454)	Control Group (N = 225)	Treatment Group (N = 229)
Child is female (%)	54.4	52.9	55.9
Child is male (%)[8]	45.6	47.1	44.1
Child age in months [mean (SD)]	59.2 (3.8)	59.3 (3.9)	59.1 (3.8)
Baseline science knowledge (*Lens on Science*) [mean (SD)]	1.37 (1.09)	1.34 (1.05)	1.39 (1.13)
Child race or ethnicity (%)			
» White	34.6	36.4	32.8
» Hispanic	18.7	18.7	18.8
» Black or African American	18.1	17.8	18.3
» Asian	3.3	2.7	3.9
» Other	2.4	1.8	3.1
» Multiracial	22.9	22.7	23.1
Child attends center care (non-K) for 30+ hours weekly (%)	25.3	22.3	28.4
Child's primary language is English (%)	75.3	76.0	74.7
Child has an IEP or 504 plan (%)	10.1	12.0	8.3
Family annual income is less than $50,000 (%)	51.1	53.8	48.4
Responding parent has high school education or lower (%)	17.2	15.5	18.8

[8] Parents reported child gender on the parent pre-survey. Response options were male, female, and other (please specify). All parents selected either male or female.

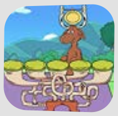
Attrition from the study was low. Among the 229 children assigned to the treatment group, 223 (97.4%) completed one or more of the post-test assessments. Rates of attrition were similar among control-assigned children, with 220 of the 225 (97.8%) control-assigned children completing at least one post-test assessment.[9] The overall attrition rate of 2.4% and differential attrition rate of 0.4% fall within the What Works Clearinghouse standards for acceptable rates of attrition for experimental studies.

Study Materials

The study team designed the study experience to be an optimized version of what young children experience in their everyday lives. We provided the participating 4- and 5-year-old children and their parents with access to the *Cat in the Hat* videos, *The Cat in the Hat Builds That* app, and printed real-world activities. Families received a tablet computer with a video player app to view the *Cat in the Hat* videos and to play games in *The Cat in the Hat Builds That*. Both apps were accessible to the children via icons on the tablet home screen. Each tablet had a data plan to enable access to the Internet, although use of the *Cat in the Hat* videos and games did not require Internet connection. *Cat in the Hat* producers organized the videos and games into six themes (Table 3, below). Resources included 28 *Cat in the Hat* videos that together totaled 3 hours and 43 minutes of content. Nine of these videos are interstitials—90-second programs designed to bridge two full-length (11-minute) video segments to fill a 30-minute television slot. Two of the five digital games in *The Cat in the Hat Builds That*, “Bridge-a-rama” and “Slidea-ma-zoo,” were designed to be adaptive and build additional levels based on past (saved) gameplay by continuously estimating player ability and choosing appropriate next levels based on factors such as demonstration of mastery and the difficulty of the levels available. In addition to the game-aligned hands-on activities embedded in the app, the study team provided families with three printed science and engineering activities from the *Cat in the Hat* website (see Appendix C for activities).

[9] Table 2 describes the full sample; a table describing the analytic sample (i.e., those who have at least one post-test assessment) is in Appendix B.

Table 3. *Cat in the Hat* Content Themes

Theme	Content Area	No. of Full-Length Videos	No. of Inter-stitials	No. of Games	No. of App-Based Hands-on Activities	Printed Real-World Parent/Child Activity
Bridge-a-rama 	Bridges	4	1	1	4	Daring Design Challenge; Measuring This and That
Slidea-ma-zoo 	Slides & friction	3	2	1	4	—
Sorta-ma-gogo 	Sorting objects	4	2	1	4	What Floats Your Boat?
Sound-a-palooza 	Sounds & soundwaves	3	—	1	4	—
Build-a-maloo 	Building & engineering	3	2	1	4	—
Be curious! 	Five senses & making observations	2	2	—	—	—

The study team and the written family guide both encouraged treatment group participants to engage with the *Cat in the Hat* digital media content for about one hour per week for eight weeks, but also suggested parents could use their own judgment to determine how much time their child spent with the resources, based on their typical media use and the child's interest. In addition to the *Cat in the Hat* materials described above, the treatment group received the following resources:

A **parent guide** with information about the study, a description of the *Cat in the Hat* materials, troubleshooting tips for using the tablet, and science-focused tips for parents that were drawn from the PBS KIDS Parents website.

A **calendar** with one of three randomized sequences of the *Cat in the Hat* content to focus on each week, including six weeks on a content theme and two weeks of free choice. The order in which families were asked to engage with the *Cat in the Hat* resources was randomized so that performance on assessments would not be influenced by how recently a child engaged with a specific set of videos, games, or activities.

Parents also received weekly text messages reminding them to engage with the target theme for the week and providing them with a link to complete a weekly media log. Participants who did not fill out their weekly media log within three days of the weekly message also received a reminder text.

Control Group. Control group participants received tablets with a data plan to enable access to the Internet. The study team members explained that they were interested in learning about how families use educational media and encouraged the control group participants to engage with educational digital media content on the tablet for one hour per week for eight weeks, or in keeping with their typical approach in guiding how much media the child uses. Third-party software blocked access to the PBS website and PBS games and video apps on the tablet, as well as to a few science-related apps that targeted content similar to *Cat in the Hat*.[10] The control group also received a parent guide with general information about the study along with troubleshooting tips for using the tablet. Control group participants received text messages reminding them about their scheduled second meeting but did not complete media logs. At the end of the eight-week study period, the study team provided control-assigned families with access to the *Cat in the Hat* resources.[11]

Table 4, below, summarizes the resources provided to families in the treatment and control groups.

[10] These apps included *Dino Train Jurassic JR., Measure Up!, PBS Parents Play and Learn Science, PEEP Trash Stash, PEEP Chirp Shapes Up, PEEP Quack's Apples, PEEP Family Science Ramps, PEEP Ciensias: Rampas, PBS KIDS Games, PBS KIDS Video and The Cat in the Hat Builds That.* Despite blocking access to the publicly available *Cat in the Hat* materials, it is possible families may have been able to access PBS KIDS shows through third-party websites such as YouTube or by using other devices.

[11] Control group families were invited to enroll in a subsequent study of the impact of a mobile messaging intervention on parent's beliefs and practices about early science learning. For more information see Registry of Efficacy and Effectiveness Studies #1627.2v1 https://sreereg.icpsr.umich.edu/framework/pdf/index.php?id=2133.

Table 4. **Resources Provided to Treatment and Control Group Participants**

	Treatment-assigned children	Control-assigned children
Technology Resources	• New tablet computer • 8 weeks of data • *The Cat in the Hat Knows a Lot About That!* Season 3 video player app • *The Cat in the Hat Builds That* games app	• New tablet computer • 8 weeks of data • *Cat in the Hat* videos and games blocked • Similar science apps, PBS KIDS website, video, and game apps blocked • Instructions for use of educational digital media content
Parent Guide Information	• Have their child use the tablet to access the *Cat in the Hat* resources • Study information and basic tips for using the tablet • Description of the *Cat in the Hat* materials and science-focused tips drawn from the *Cat in the Hat* website	• Have their child use the tablet to access educational media of their choosing • Study information and basic tips for using the tablet
Text Messages	• Weekly text message reminders about the *Cat in the Hat* content of the week, media log link • Text message reminders of data collection appointments	• Text message reminders of data collection appointments
Other Resources	• Calendar indicating each week's *Cat in the Hat* focal theme	—

Data

The study team collected three types of data: outcomes measures, parent measures, and tablet usage data.

Outcomes Measures

The study team assessed children's science and engineering knowledge and science practices using two types of assessments:

1. a modified version of Lens on Science, an externally developed measure of young children's physical science and engineering knowledge, and

2. researcher-developed measures, the Hands-On Preschool Assessments of Physical Science and Engineering.

» **Lens on Science.** The computer-administered, adaptive, item response theory-based Lens on Science (Lens) assessment for preschool children was developed by Daryl Greenfield and colleagues at the University of Miami (Greenfield, 2015). The Lens on Science is aligned with the Framework for K-12 Science Education (National Research Council, 2012) and intended to assess young children's knowledge of core ideas in four science disciplinary areas (life science, earth/space science, physical science, and engineering and technology); science practice skills; and crosscutting concepts. It was designed to detect growth in learning for preschool children from low-income households. It was chosen as an outcome measure for this study because of the psychometric evidence of its validity and reliability. The assessment takes approximately 20–25 minutes to administer, including a 7- to 8-minute screener that ensures a child understands and knows how to respond to its three different item types. During the assessment itself, children are given a computer adaptive set of items matched to their abilities based on a pool of 498 items.

- At pre-test, all children completed the full adaptive Lens assessment, which served as a baseline for later analyses to ensure that both groups of children had similar science knowledge at the start of the study.

- At post-test, all children completed a subset of items from the full Lens (Lens-Modified) that covered topics in the physical science and engineering and technology domains. Details on the post-test are in Appendix D.

» **Hands-On Preschool Assessments of Physical Science and Engineering (Performance-Based Assessments).** The research team developed three performance-based tasks to assess children's understanding of the role of material properties (strength and length) and forces in structural stability, the role of material properties (texture) and forces (friction) on movement down an incline, and how objects can be sorted based on their material properties and uses. The tasks also emphasize the science practices that are a focus of *Cat in the Hat*, as well as the crosscutting concepts that unite the core ideas of science and engineering related to stability and change, structure and function, cause and effect, and systems. These tasks were developed to complement the Lens and addressed a narrower subset of content covered in *Cat in the Hat*. Children completed these tasks at post-test. In order to limit the total assessment time to no more than 35 minutes, each child was randomly assigned two of the three assessments. Both the specific performance tasks and the order in which the tasks were delivered were determined randomly before the initiation of the study. Additional details about each of these assessments and their scoring is in Appendix D.

- **Length, Strength, and Stability.** This task assesses a child's understanding of how the properties of objects (such as size and shape) and materials (such as hardness and flexibility) make them suitable for a purpose. Children are provided with a group of objects of different lengths and strengths and asked to choose the most suitable object for building a bridge that can support weight.
- **Surfaces and Friction.** This task assesses a child's understanding of how the properties of materials and forces—friction in particular—influence the motion of objects. Children are provided with three slides with differently textured (rough/smooth/sticky) surfaces and asked to choose the slide that will enable a toy figure to slide down the fastest.
- **Colors, Shapes, and Uses.** This task assesses a child's understanding that different objects can be described and categorized based on their observable properties and on how they are used. Children are provided with different sets of objects and are asked to sort them based on color, shape, and standard use (for eating, making art, and playing).

Parent Measures

» **Parent Survey.** Parents completed pre- and post-surveys via the survey platform Qualtrics. All surveys were written at approximately a fifth-grade reading level and were available in both English and Spanish.

- **Parent Pre-survey.** This included demographic questions regarding parent and child backgrounds. The survey also asked about parent confidence related to science and technology, the child's use of technology and media, and the child's prior exposure to PBS children's media properties, including *Cat in the Hat*. Parents completed the pre-survey before assignment to condition.
- **Parent Post-survey.** This included questions regarding parent report of child's engagement in science activities and use of the *Cat in the Hat* science vocabulary over the last month, such as "cause," "effect," "measure," "build," and "vibration," as well as science vocabulary not related to the *Cat in the Hat* resources, such as "plant," "gravity," and "chemical." The survey also asked treatment-assigned

families questions about parent perceptions of the impact of *Cat in the Hat* on their child's learning and development. Treatment- and control-assigned families completed the survey at the conclusion of this study (Week 8).

» **Media Log.** The media log was a brief (5-minute) survey with several questions about the child's engagement in the study materials over the prior week. Parents in the treatment group received one text message per week that reminded them of the content of focus for the week and included a link to complete the weekly media log. Participants who did not complete the weekly media log within a few days received an additional text reminder to do so.

» **Parent Interviews.** Four parents at each of the five sites (20 total, approximately 5% of the overall sample) were sampled to participate in a brief (10- to 15-minute) phone interview during the week after their post-test data collection. Parents were sorted within sites into one of four groups based on their child's use of the *Cat in the Hat* resources as measured by the tablet usage data (low, medium low, medium high, or high). The study team randomly selected one parent from each usage group within each site. Interviewers asked parents to describe their child's overall experience using the *Cat in the Hat* videos, games, and hands-on experiences during the study. Parents were compensated $25 for their time. Interviews were conducted in either English or Spanish.

Tablet Usage Data

The study team worked with a third-party app developer to develop the *Cat in the Hat* video player app, which was how children watched the *Cat in the Hat* videos. The app logged which the *Cat in the Hat* videos each treatment-assigned child watched and for how many minutes. In addition, PBS developed a research version of *The Cat in the Hat Builds That*, identical to the publicly available app except that it logged the games each treatment-assigned child played and their use time. Finally, the study team installed

third-party data tracking software, InterGuard, on both treatment and control tablet computers. The software collected usage data in two ways: **(1)** the name of each online web page that was accessed using the device, along with the date, and **(2)** the name of any apps that were opened and the time of use. The primary purpose of the third-party software was to monitor the control group for possible use of the *Cat in the Hat* resources. This software also blocked control families from accessing the *Cat in the Hat*-related resources. Data were collected using participant and tablet identification numbers to protect children's identities. We created summaries of the tablet usage data using frequencies, means, and standard deviations to understand the breadth and depth of children's engagement with the videos and games and their engagement across the weeks of the study among children in the *Cat in the Hat* group.[12]

For more information on data collection instruments and analysis procedures, see Appendices A and B.

Data Analysis

We used multivariate linear regression (ordinary least squares) to examine differences in post-test outcomes between the treatment- and control-assigned participants.[13] For the outcome models—child assessments and parent-reported measures— all regression models included controls for children's baseline scores on the Lens on Science assessment. This approach permitted us to identify the impact of *Cat in the Hat* on children's understanding of science and engineering concepts and practices while controlling for any differences in their initial skills. Children who were missing records or had incomplete records of either their initial skills (Lens pre-test) or post-test data were not included in the analyses. We tested the sensitivity of the regression analyses by including relevant child and family demographic characteristics.[14] The results of the analyses for research questions 1–4 can be interpreted as supporting causal inferences about the impact of *Cat in the Hat* on children's science and engineering knowledge and practices. For these analyses, we characterize the magnitude of the effect size using benchmarks drawn from Kraft's (2019) research on the distribution of effect sizes from a large sample of randomized studies using standardized assessments as educational outcomes. This research characterizes impacts of less than 0.05 SD as small, effects of 0.05 SD to less than 0.20 SD as medium, and 0.20 SD or greater as large.[15]

We also conducted exploratory analyses on differences in the impact of the *Cat in the Hat* resources for subgroups of children and families (research question 7). We conducted these analyses by adding interactions of the subgroup and treatment indicators to the multivariate regression analyses used to answer research questions 1–4.

We explored how the amount of time that a treatment-assigned child used the *Cat in the Hat* resources was associated with the child's learning during the study. We did this by conducting regression analyses of the relationship between usage metrics and the residual value produced by the impact regressions on the four primary outcome variables (described above). These residual values represent the difference between a child's observed value for the outcome variable

[12] Families were instructed that only the child selected to participate in the study should use the tablet during the study. However, it is possible that other members of a household, such as siblings, may have used the *Cat in the Hat* materials during the study.

[13] Appendix B provides additional detail on our data analyses.

[14] We included the relevant family and child demographic characteristics in our analyses We selected these child and family characteristics for inclusion in the analyses if they were statistically significant predictors of the outcomes in bivariate analyses. See Appendix B for model-specific covariates.

[15] Note that these classifications of effect size magnitude differ from those suggested by Cohen (1969): <0.2 Small, 02.- 0.5 Medium, >0.8 Large).

and the value that was predicted from the impact regression analyses on the respective outcome. For each outcome variable, we conducted separate regression analyses of the relationship between the residual and the number of minutes of video usage, the number of minutes of game usage, and the percentage of the available games or videos the child accessed at least once over the course of the study.[16] For analyses with the Lens data, we used usage metrics that included use of all the *Cat in the Hat* resources. For the analyses using the performance-based assessments, we used information on the usage of content related to those learning goals. For example, in examining the relationship between the residual score on the Length, Strength, and Stability task and usage, we examined the amount of time children spent watching videos that were part of the "Bridge-a-rama" content theme only. This approach provided information about whether the amount of usage of the *Cat in the Hat* resources was related to whether a child performed better or worse than expected based on his or her science and engineering skills at preschool entry and demographic characteristics.

These analyses were exploratory and did not support causal inferences about the impact of usage on children's learning.

For the quantitative data regarding family background and engagement with the *Cat in the Hat* resources collected through parent surveys and tablet usage data, the study team conducted descriptive analyses by calculating basic frequencies, percentages, means, and standard deviations.

The study team analyzed parent interview data to identify common themes about how families engaged with the *Cat in the Hat* materials using four categories: use and engagement, challenges experienced, perceived benefits for children, and perceived benefits for parents. After identifying subthemes that emerged in each of the categories, the study team selected key quotes to illustrate each of the themes.

[16] We report associations that are statistically significant at $p < .03$. We arrived at $p < .03$ by using an alpha level of $p < .10$ and accounting for multiple comparisons using a Bonferroni correction.


sembler

Results

The analyses indicate that engagement with *Cat in the Hat* resources had medium to large **positive impacts** on children's learning of specific science and engineering concepts and practices.

Impacts of the *Cat in the Hat* Resources on Children's Science Knowledge and Practices

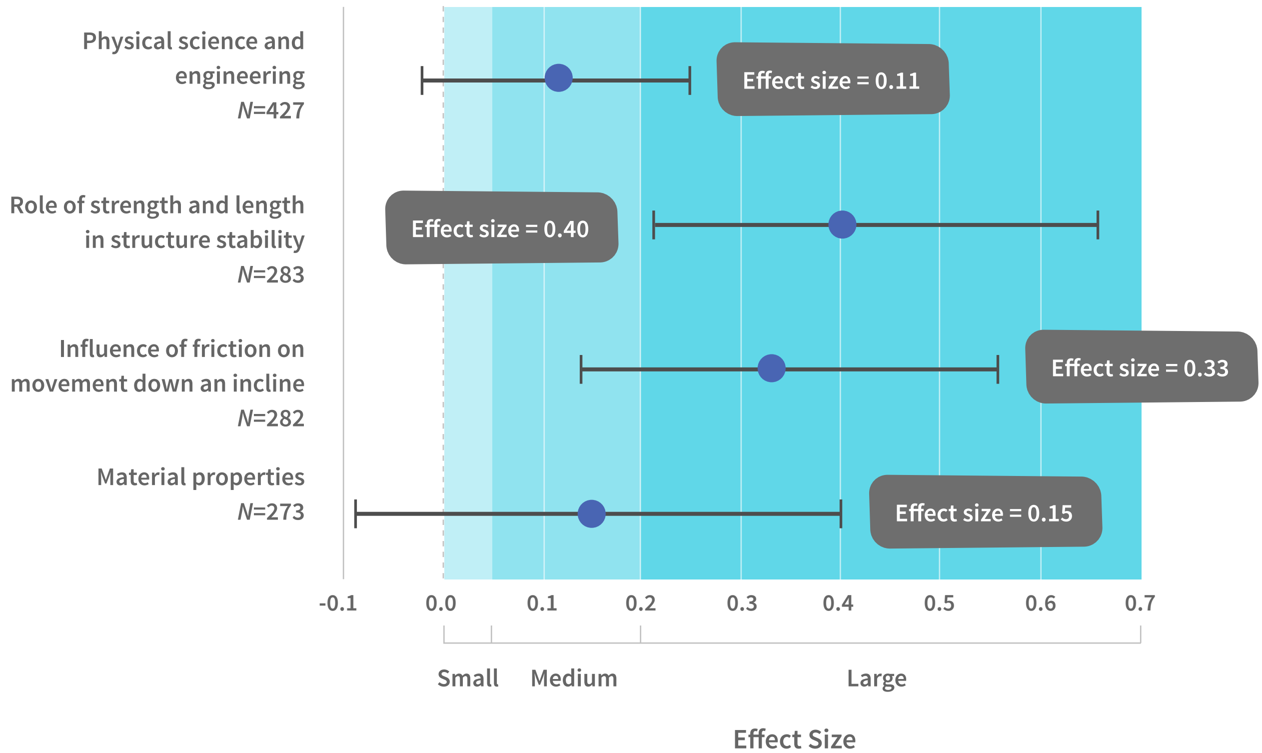
The analyses indicate that access to *Cat in the Hat* resources had medium to large positive impacts on children's learning of specific science and engineering concepts and practices. Analyses provide promising evidence that *Cat in the Hat* also supported improvements in broader physical science and engineering knowledge. We describe the magnitude of these impacts using Cohen's *d* effect sizes (represented as "*d*" below).

Specifically, the results indicate that providing children with eight weeks of access to the *Cat in the Hat* resources substantively improved

- children's understanding of how the properties of objects and materials (strength and length) and natural and applied forces contribute to the stability of structures ($d = 0.40$, $p < .001$); and
- children's understanding of how the properties of materials (texture) and forces (friction) influence how objects move ($d = 0.38$, $p < .01$).

We also found suggestive evidence of a small positive impact on children's understanding of material properties (ability to sort objects by size, color, shape, and use). However, this impact was not statistically significant ($d = 0.15$, $p = .18$). Similarly, we found suggestive evidence that the *Cat in the Hat* resources improved children's broader physical science and engineering knowledge and practices as measured by a subset of items from the Lens on Science (d = 0.11, p = .12). Although this impact is not statistically significant at conventional levels, the relatively low probability that this finding is an error (that is, 12% likelihood versus the 5%-10% convention) provides promising evidence of the *Cat in the Hat's* impacts on children's broader understanding of physical science and engineering knowledge and practices. These results are displayed graphically in Figure 1.

Figure 1. *Cat in the Hat* resources' impacts on children's science and engineering knowledge and practices



Note: All models include controls for baseline score on the Lens on Science. Effect size refers to the standardized mean difference between two groups in standard deviation units (Cohen's d). Characterizations of effects as small, medium, and large are based on Kraft (2019). Dots represent effect size from the regression models. Bars represent the 95 percent confidence intervals around the effect size.

Impacts of *Cat in the Hat* Resources on Parent Reports of Science and Engineering Engagement

We conducted exploratory analyses to better understand impacts of the *Cat in the Hat* resources on children's engagement with science and use of science-related vocabulary. Results from parent reports suggest that providing children with access to the *Cat in the Hat* resources increased their excitement about science (Figure 2, page 29). On average, parents in the treatment group rated their children's excitement about science higher than did parents in the control group ($d = 0.24$, $p < .05$). There were no differences between treatment group parents' reports of children's excitement about engineering. Treatment-assigned parents also reported that children engaged in more science activities over the past month than did control-assigned parents ($d = 0.20$, $p < .05$).

The exploratory analyses suggest that the *Cat in the Hat* resources increased children's use of program-related vocabulary as reported by their parents, such as cause, effect, measure, build, and vibration. Treatment-assigned parents reported that their children used more study-related science vocabulary in the past month than control-assigned families reported ($d = 0.18$, $p < .10$). These results should be interpreted with caution as they are no longer statistically significant when we control for student and family characteristics. We did not find treatment-control differences in parent reports of vocabulary words not related to the *Cat in the Hat* resources.

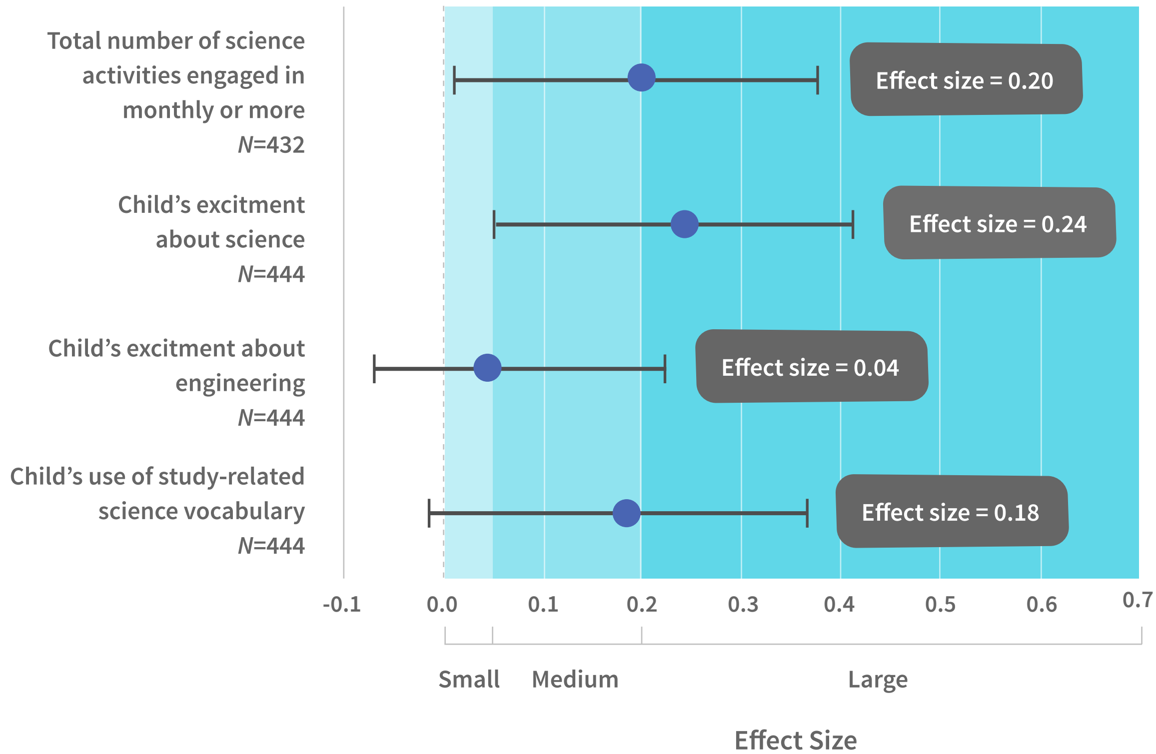
What Parents Say

"She always likes to build and everything. But now it's more not just building but seeing the differences in the buildings she makes. Is one stronger or not? She's questioning—why one fell and the other one didn't. So she's looking at the structure of the buildings she makes." *—Parent*

"I think there was more of a deliberateness or a consciousness in me to explain the little things that might just have passed by if it weren't for the app. Like, for the simple concept of gravity, I did get to discuss that with him because when we were looking at the slide and how to build it, how inclined it should be, and things like that." *—Parent*

"She was talking about different cause and effects, which I think is one of the videos, and like how one thing causes another thing to happen. She was talking about that later and about building bridges and how different weight is supported by different materials." *—Parent*

Figure 2. ***Cat in the Hat*** **impacts on parent-reported outcomes**

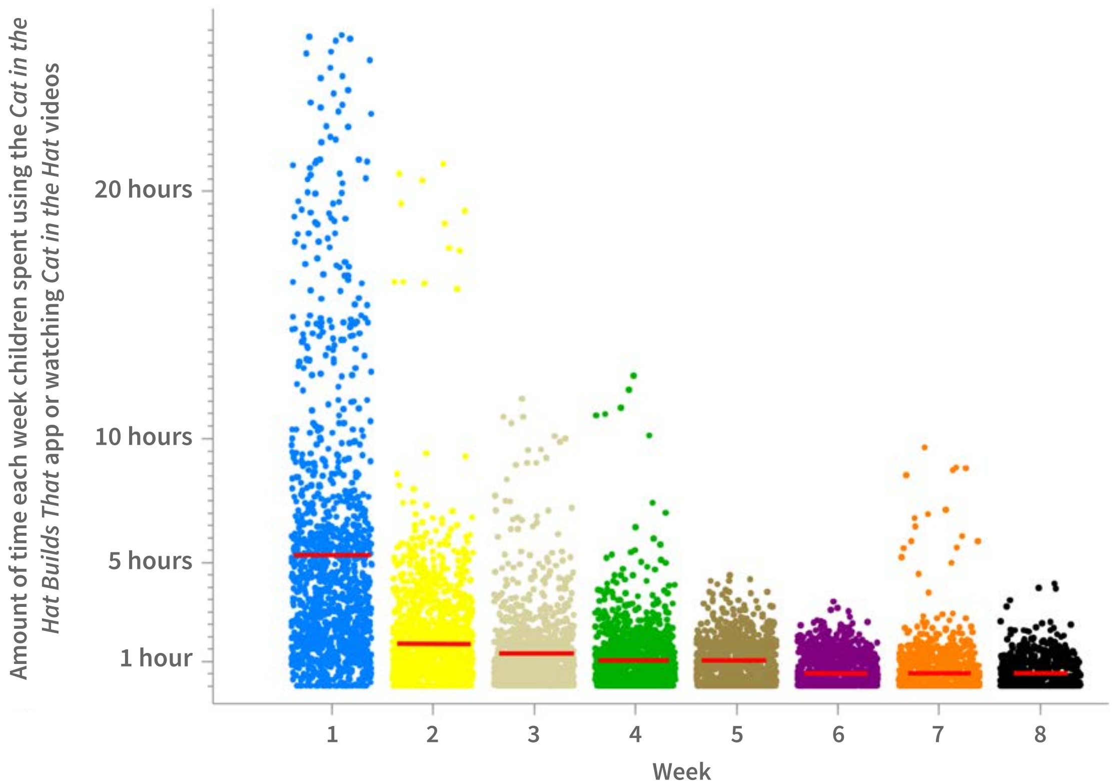


Note: All models include controls for baseline score on the Lens on Science. Effect size refers to the standardized mean difference between two groups in standard deviation units (Cohen's *d*). Characterizations of effects as small, medium, and large are based on Kraft (2019). Dots represent effect size from the regression models. Bars represent the 95 percent confidence intervals around the effect size.

Use of *Cat in the Hat* Resources

Over the eight-week study period, children in the treatment group on average played the *Cat in the Hat* digital games for a total of 2 hours 48 minutes and watched the *Cat in the Hat* videos for 4 hours 20 minutes. Video and game usage differed substantially across children: While approximately one quarter (24%) of treatment-assigned children engaged with the videos or games for less than 2 hours across the eight weeks, over half (59%) of the children engaged with these materials for more than 20 hours. Video and game usage declined sharply after the first week of the study. During Week 1, children engaged with the study videos and games for an average of 5 hours 14 minutes. Average usage of the *Cat in the Hat* games and videos in the second week of the study dropped to just under 2 hours, and then declined to less than half an hour by Week 8.

Figure 3. **Minutes spent using the *Cat in the Hat* videos and games by week**

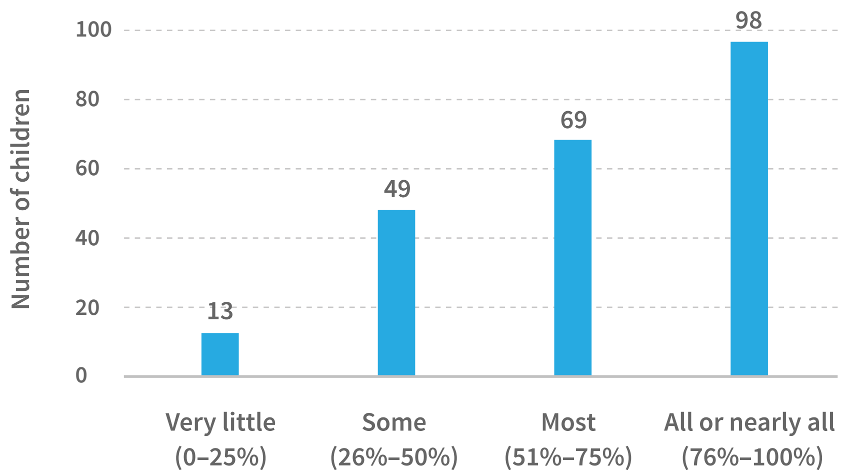


Mean	5 hrs, 14 mins	1 hr, 56 mins	1 hr, 16 mins	1 hr, 2 mins	51 mins	32 mins	34 mins	27 mins

Note. Each dot represents the amount of time a single child spent using the *Cat in the Hat Builds That* app or watching the *Cat in the Hat* videos in a given week. Red horizontal bars represent the average amount of time for all children in that week.

→ Three quarters of the children in the treatment group (75%) accessed materials from four of the six *Cat in the Hat* content themes for 20 minutes or more across the eight-week study, and nearly half of the children (47%) accessed material from all six content themes.

Figure 4. **Percentage of all *Cat in the Hat* games and apps used by treatment group children (*N*=229)**



Analyses of tablet usage data suggest that about half (49%) of the children accessed at least one of the real-world activities embedded in *The Cat in the Hat Builds That*. Children accessed the hands-on activities 1.4 times on average, for an average time of under one minute (0.48 minutes).

Among children in the control group, six were able to circumvent the app-blocking software, and one of these children accessed more than 20 minutes of *Cat in the Hat* content over the study period.

Parent responses during interviews indicate that children enjoyed playing the games and watching the videos. Parents reported that the "Bridge-a-rama" and "Slidea-ma-zoo" games were particularly popular with children. In addition to using the *Cat in the Hat* apps, children also watched the *Cat in the Hat* videos on other platforms such as Netflix and YouTube.

Parents also reported that their children enjoyed the real-world activities, although they used these activities less regularly. Interview data suggest that the resources also fostered more informal everyday science and engineering activities. Close to three quarters (73%) of parents reported in their weekly media logs that they used a *Cat in the Hat* hands-on activity with their child at least once during the study period. Interviewed parents indicated that the videos and games were a great source of ideas for everyday activities and that their children mimicked activities they saw in the media with their toys, such as building bridges or making patterns. Some interviewed parents also noted that the resources gave them ideas for science activities to do with their children at home and how to incorporate science into everyday activities in a way that was fun and child-friendly. For example, parents reported becoming more aware of their own language and trying to include more science talk and scientific concepts and ask more questions.

When asked about study challenges, five parents reported that they had trouble getting their child to follow the suggested schedule because their child wanted to choose what to do or repeat an activity rather than do the ones in the suggested calendar, or that their child stopped playing games after a short period because they were not challenging enough or did not have enough levels—perhaps a reason for the decline in usage of the *Cat in the Hat* videos and games over the course of the study (see Figure 3). Some parents referenced their own lack of knowledge about science concepts and about how to engage children in science.

The majority of interviewed parents indicated that the activities helped their child learn about a variety of science concepts and skills. Specifically, parents talked about an increase in skills related to asking questions, predicting, problem solving, trying different methods, and testing, as well as critical thinking and curiosity. Parents also described learning related to concepts such as math and measurement; the properties of different materials; and content related to plants, insects, and sounds. Parents did not use the term "engineering" in describing learning, but many referenced "building." A few parents also noted enhanced social-emotional development, including patience, confidence, and independence. Finally, many parents felt that children's interest in science and engineering increased or was enhanced, although many reported that their children were already interested in science before the study.

What Parents Say

"…we did the one with the colored ice. We colored the water, then we froze it, and then we watched it melt. But we also did other variations, like we tried melting it on top of a piece of paper and on top of a piece of table napkin, just to see also how the colors will be absorbed. We then just added to the experiment." *—Parent*

Relationship between use of the *Cat in the Hat* resources and performance on science and engineering measures.

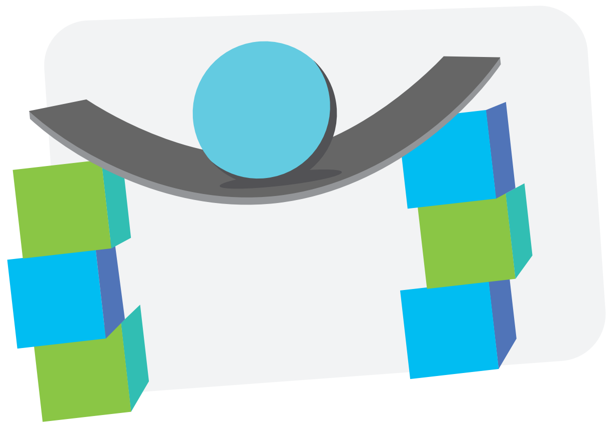
In general, we did not identify a consistent relationship between the amount of time children used the *Cat in the Hat* games or watched the *Cat in the Hat* videos and their performance on post-test measures. We did find some evidence to suggest that the more time children watched the "Bridge-a-rama," videos, the more they outperformed our predictions on the Length, Strength, and Stability assessment. Similarly, the more time children spent playing the "Slidea-ma-zoo" game, the more they outperformed predictions on the Surfaces and Friction task. However, the positive relationship between use and outcomes appears to weaken at very high levels of use (approximately 2 hours of "Bridge-a-rama" video use and 3 hours 30 minutes of "Slidea-ma-zoo" gameplay). Future analyses will explore this relationship in more depth. We did not observe any statistically significant relationships between usage variables and children's performance on the Lens or the Colors, Shapes, and Uses measures.

Impacts of *Cat in the Hat* by Subgroup

We examined whether the impact of *Cat in the Hat* on children's science knowledge and practice differed depending on children's prior science knowledge and demographic characteristics. Specifically, we considered whether the impact of providing access to the *Cat in the Hat* resources differed depending on children's baseline Lens scores, age, gender, ethnicity (White versus non-White); whether the parent had a high school education or higher; and whether the child lived in an English-only household, attended center-based care for 30 hours or more, and was in kindergarten.[17]

We did not find evidence that the impact of *Cat in the Hat* on children's physical science and engineering knowledge, as measured by the Lens on Science assessment, was statistically significantly different for children in any of the subgroups described. Similarly, we did not find evidence that the impact of the *Cat in the Hat* resources on children's understanding of the role of material properties (strength and length) and forces in structural stability or the role of material properties (texture) and forces (friction) on movement down an incline varied depending on their baseline physical science and engineering knowledge or background characteristics.

However, we did find that the impacts of *Cat in the Hat* on children's understanding of how objects can be sorted based on their material properties and uses were stronger for children whose families spoke English only at home than for children whose families spoke a language other than English at home ($p < .01$). These findings suggest that the *Cat in the Hat* resources differentially affected children's understanding of material properties as measured by the Colors, Shapes, and Uses performance-based assessment: Children who spoke English only at home seemed to benefit more from access to the *Cat in the Hat* resources relative to their understanding of material properties than did children from families who spoke languages in addition to or other than English at home.



[17] We report associations that are statistically significant at $p < .0125$. We arrived at $p < .0125$ by using an alpha level of $p < .10$ and accounting for multiple comparisons using a Bonferroni correction.


FAMILY SIZE
Cinnamon Toast Crunch
mmm.....REAL cinnamon

Discussion

The results of this study indicate that providing young children and their families with **access to science-based media resources** can support their science- and engineering-related learning.

Specifically, our study found that providing families with access to the *Cat in the Hat* videos and digital games improved children's understanding of physical science concepts related to the two physical science core ideas of matter and forces: the role of material properties (strength and length) and forces in structural stability, and the role of texture on friction and movement down an incline. Children who were provided access to the *Cat in the Hat* resources outperformed those who did not have access on measures of how the properties of objects (such as size and shape) and materials (such as hardness and flexibility) make them suitable for different purposes, and how different forces (pushes and pulls) can cause objects to move and influence the stability of a bridge. Children with *Cat in the Hat* access also learned more about how the properties of objects, materials, and forces (including friction) influence the motion of objects. We also found promising evidence that children's understanding of sorting and their understanding that different objects can be described and categorized based on their observable properties or functions improved as a result of using the *Cat in the Hat* resources. Although not statistically significant, and thus less certain, results from the externally developed assessment of children's physical science and engineering knowledge and practices also provided suggestive evidence of the *Cat in the Hat's* impact on broader science and engineering learning.

The treatment-control differences in children's performance on the performance-based assessments also provided evidence of children's ability to transfer learning from a digital environment to a real-world setting. This evidence is notable in the context of a substantial body of research that has documented the challenges young children face in transferring knowledge from videos because of difficulties in separating fantasy from reality (Woolley & Ghossainy, 2013), confusion about how videos and games relate to the real world (Mares & Sivakumar, 2014), and difficulties in transferring two-dimensional experiences to a three-dimensional world (Anderson & Pempek, 2005). The results of this study offer evidence that fantasy-based storytelling can support young children's accurate scientific understanding if the narrative content focuses on factual science and engineering concepts and practices.

Among the strengths of this study was the attention to minimizing bias wherever possible. Attrition was minimal and met the What Works Clearinghouse standards. Study team assessors and analysts remained blind to the participants' treatment or control status throughout the study. Parents in the control condition were not told they were control families. These processes allow us to feel confident in making casual inferences regarding the impacts of the *Cat in the Hat* resources.

Young children spend substantial time using digital media. A critical question raised by these practices is whether this time spent engaging with videos and digital games can be productive. The increasing ubiquity of media in our daily lives and the opportunity that media offer to scale up an educational intervention at a low cost add urgency to this question. A growing body of research on media resources suggests that certain kinds of media experiences can help children learn literacy, math, and socio-emotional skills (Anderson et al., 2001; Fisch, Truglio, & Cole, 1999; Hirsh-Pasek et al., 2015; Hurwitz, 2018; Kirkorian, Wartella, & Anderson, 2008; Schmidt & Anderson, 2007). However, very little is known about how and whether media can help young children learn science concepts and practices, and few science-based media resources are available for young children. This study adds new knowledge about the promise of media to support science and engineering learning. Moreover, despite the research evidence about the difficulty children face in transferring knowledge to new contexts, this study provides important findings about the possibilities media resources hold for helping children develop knowledge that can transfer from a digital to a physical environment.

The size of the impacts of *Cat in the Hat* on learning is particularly notable given the relatively low per-child cost to scale up access to the *Cat in the Hat* resources compared to typical effective science interventions such as resource-intensive science kits or professional development (Slavin, Lake, Hanley, & Thurston, 2014). In this context, the size of the *Cat in the Hat* resource impacts on learning (ranging from 0.11 to 0.40 SD) are substantial.

The results of the study also raise new questions. The study focused on a relatively narrow set of physical science and engineering concepts and practices; future research should examine whether media hold promise for other concepts in science that are more abstract. In addition, measures of parent engagement in the media in this study relied almost entirely on self-report; an important question related to children's ability to transfer learning from a digital environment to the physical world is the extent to which parents facilitated this transfer through, for example, real-world activities and talking about the *Cat in the Hat* content.

Limitations

This study used a rigorous experimental design. Assignment and data collection procedures prevented assessors and families from knowing their treatment condition, and attrition from the study was minimal. We nevertheless encourage readers to consider the following limitations when interpreting study findings.

Generalizability of Study Sample. Study sites were selected to include multiple geographic regions (Northeast, Midwest, Southwest, and West), but site selection was not designed to produce a nationally representative sample. Further, participating families were volunteers recruited through social media posts regarding participation in a digital media study. They and their children may thus be different from parents who do not use social media or are not interested in participating in a digital media study.

Generalizability of Study Experiences. The *Cat in the Hat* materials are widely available through the PBS KIDS website and local public media stations, but the ways study children accessed them differed in important ways from how they would be accessed by children outside of the study. First, all digital materials (games and videos) were preloaded onto tablet computers and were accessible to children via icons on the tablet home screen. This kind of accessible display most likely encouraged treatment-assigned children to access the materials more often than they would have otherwise, particularly given that episodes are not available all at once—rather, several episodes at a time are broadcast over a year. Second, all participating children were given a new tablet computer as part of the study. Providing this resource does not limit the internal validity of the study findings as both treatment- and control-assigned children were given the new tablets. It is nevertheless likely that the novelty of the new device encouraged treatment-assigned children to use the *Cat in the Hat* resources more than they would have without a new device. Third, treatment-assigned parents were encouraged to support their children's use of the *Cat in the Hat* materials through reminders about study material. Again, these reminders most likely led children to use the materials more often than they would have outside the study.

Study Measures. We used a combination of existing measures and measures that we adapted or created for use in this study. The study team designed these measures to reflect children's understanding of the knowledge and practices that are the targets of the *Cat in the Hat* content. We do not have information on the validity and reliability of these measures. A detailed description of our measure development procedures is in Appendix D.

Implications

We see three **key implications** from this study.

1 **First, interacting with the Cat in the Hat resources measurably improved some aspects of children's knowledge of science and science and engineering practices.**

Although our certainty about the impacts is stronger for a subset of content knowledge and practices in physical science and engineering, the results provide consistent evidence that *Cat in the Hat* helps children learn science and engineering concepts and practices.

2 **Second, the strong results on the friction and incline and the structure and stability performance-based measures indicate that children's experiences manipulating materials in a digital context can transfer to mastery of those practices and knowledge in the physical world.**

This finding provides a substantial contribution to overall research literature on how children learn from digital media.

Recent research on effect sizes and educational interventions (Kraft, 2019) calls for considering the magnitude of a program's impact within the broader context of the costs and risks associated with the interventions. Given both the easy scalability and low per-child cost of providing access to the

Cat in the Hat resources, the positive results from this study also appear meaningful in the context of other typically resource-heavy science interventions. In other words, we should interpret these findings differently than we would those for a classroom science intervention over the course of a school year that requires training of a teacher and other expensive resources and that is challenging to scale up. The cost and effort to implement this intervention is less, and so a smaller impact on learning is more meaningful. For example, before the eight-week *Cat in the Hat* experience, children were most likely already engaged in similar activities—using media—but possibly with a less-educational resource, given the general lack of quality of educational apps (Hirsh-Pasek et al., 2015). In other words, even in the face of medium to small effect sizes, simply directing families to high-quality media and to some guidance for parents about how to use the media seems like an important step in improving young children's engagement in science.

3 Third, these findings also point to the importance of careful design that aligns with key learning goals relevant for young children.

The *Cat in the Hat* media were developed with an intensive focus on core science and engineering concepts and practices and with the engagement of content experts and iterative formative development processes, all of which likely contributed to the resulting high-quality media. The positive results of this study demonstrate that well-designed, high-quality, and learning-focused content can support learning outcomes among young children.

References

Alexander, J. M., Johnson, K. E., & Kelley, K. (2012). Longitudinal analysis of the relations between opportunities to learn about science and the development of interests related to science. *Science Education, 96*(5), 763–786.

Anderson, D. R., Huston, A. C., Schmitt, K. L., Linebarger, D. L., Wright, J. C., & Larson, R. (2001). Early childhood television viewing and adolescent behavior: The recontact study. *Monographs of the Society for Research in Child Development, 66*(1), 1–154.

Anderson, D. R., & Pempek, T. A. (2005). Television and very young children. *American Behavioral Scientist, 48*(5), 505–522.

Bandura, A. (1965). Influence of models' reinforcement contingencies on the acquisition of imitative responses. *Journal of Personality and Social Psychology, 1*(6), 589–595.

Bonus, J. A., & Mares, M. L. (2019). Learned and remembered but rejected: Preschoolers' reality judgments and transfer from *Sesame Street. Communication Research, 46*(3), 375–400.

Brenneman, K., Brown, B., Sweetman, S., Worth, K., CAST, Conn-Powers, M., & Retamal, F. (2018). PBS KIDS science learning framework. Crystal City, VA: PBS KIDS.

Bustamante, A. S., White, L. J., & Greenfield, D. B. (2017). Approaches to learning and school readiness in Head Start: Applications to preschool science. *Learning and Individual Differences, 56*, 112–118.

Fisch, S. M., Truglio, R. T., & Cole, C. F. (1999). The impact of Sesame Street on preschool children: A review and synthesis of 30 years' research. *Media Psychology, 1*(2), 165–190.

French, L. (2004). Science as the center of a coherent, integrated early childhood curriculum. *Early Childhood Research Quarterly, 19*(1), 138–149.

Fusaro, M., & Smith, M. C. (2018). Preschoolers' inquisitiveness and science-relevant problem solving. *Early Childhood Research Quarterly, 42*, 119–127.

Gerde, H. K., Schachter, R. E., & Wasik, B. A. (2013). Using the scientific method to guide learning: An integrated approach to early childhood curriculum. *Early Childhood Education Journal, 41*(5), 315–323.

Greenfield, D. B., Jirout, J., Dominguez, X., Greenberg, A., Maier, M., & Fuccillo, J. (2009). Science in the preschool classroom: A programmatic research agenda to Improve science readiness. *Early Education and Development, 20*(2), 238–264.

Greenfield, D. B. (2015). Assessment in early childhood science education. In K. Trundle, & M. Sackes (Eds.), *Research in early childhood science education* (pp.353–380). New York, NY: Springer Publishing.

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.

Hurwitz, L. B. (2018). Getting a read on Ready To Learn media: A meta-analytic review of effects on literacy. *Child Development, 90*(5), 1754–1771.

Kraft, M. (2019). *Interpreting effect sizes of education interventions.* (EdWorkingPaper: 19-10). Retrieved from Annenberg Institute at Brown University website: https://doi.org/10.26300/8pjp-2z74.

Kearney, M. S., & Levine, P. B. (2019). Early childhood education by television: Lessons from *Sesame Street. American Economic Journal: Applied Economics, 11*(1), 318–50.

Kirkorian, H. L., Wartella, E. A., & Anderson, D. R. (2008). Media and young children's learning. *The Future of Children, 18*(1), 39–61.

Kuhn, D. (2010). What is scientific thinking and how does it develop? In U. Goswami (Ed.), *Wiley-Blackwell handbook of childhood cognitive development* (2nd ed., pp. 497–523). Oxford, UK: Blackwell Publishing Ltd.

Lauricella, A. R., Gola, A. A. H., & Calvert, S. L. (2011). Toddlers' learning from socially meaningful video characters. *Media Psychology, 14*(2), 216–232.

Linebarger, D. N., Brey, E., Fenstermacher, S., & Barr, R. (2017). What makes preschool educational television educational? A content analysis of literacy, language-promoting, and prosocial preschool programming. In R. Barr & D. Linebarger (Eds.), *Media exposure during infancy and early childhood* (pp. 97–133). Cham, CH: Springer International Publishing AG.

Lu, A. S., Buday, R., Thompson, D., & Baranowski, T. (2016). What type of narrative do children prefer in active video games? An exploratory study of cognitive and emotional responses. In S. Y. Tettegah & D. H. Wenhao (Eds.), *Emotions, technology, and digital games* (pp. 137–155). San Diego, CA: Academic Press.

Mares, M. L., & Pan, Z. (2013). Effects of Sesame Street: A meta-analysis of children's learning in 15 countries. *Journal of Applied Developmental Psychology, 34*(3), 140–151.

Mares, M. L., & Sivakumar, G. (2014). "Vámonos means go, but that's made up for the show": Reality confusions and learning from educational TV. *Developmental Psychology, 50*(11), 2498.

National Research Council. (2012). *A framework for K–12 science education: Practices, crosscutting concepts, and core ideas.* Washington, DC: The National Academies Press.

National Research Council. (2013). *Next Generation Science Standards: For states, by states*. Washington, DC: The National Academies Press.

Nayfeld, I., Fuccillo, J., & Greenfield, D. B. (2013). Executive functions in early learning: Extending the relationship between executive functions and school readiness to science. *Learning and Individual Differences, 26*, 81–88.

Patrick, H., Mantzicopoulos, P., & Samarapungavan, A. (2009). Motivation for learning science in kindergarten: Is there a gender gap and does integrated inquiry and literacy instruction make a difference. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching, 46*(2), 166–191.

Peterson, S. M., & French, L. (2008). Supporting young children's explanations through inquiry science in preschool. *Early Childhood Research Quarterly, 23*(3), 395–408.

Piasta, S. B., Pelatti, C. Y., & Miller, H. L. (2014). Mathematics and science learning opportunities in preschool classrooms. *Early Education and Development, 25*(4), 445–468.

Rasmussen, E. E., Shafer, A., Colwell, M. J., White, S., Punyanunt-Carter, N., Densley, R. L., & Wright, H. (2016). Relation between active mediation, exposure to Daniel Tiger's Neighborhood, and US preschoolers' social and emotional development. *Journal of Children and Media, 10*(4), 443–461.

Rideout, V. (2017). *The common sense census: Media use by kids age zero to eight*. San Francisco, CA: Common Sense Media.

Rutten, N., Van Joolingen, W. R., & Van der Veen, J. T. (2012). The learning effects of computer simulations in science education. *Computers & Education, 58*(1), 136–153.

Schmidt, M. E., & Anderson, D. R. (2007). The impact of television on cognitive development and educational achievement. In N. Pecora, J. P. Murray, & E. A. Wartella (Eds.), *Children and television: Fifty years of research* (pp. 65–84). Mahwah, NJ: Lawrence Erlbaum Associates.

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.

Silander, M., Grindal, T., Hupert, N., Garcia, E., Anderson, K., Vahey, P. & Pasnik, S. (2018). *What parents talk about when they talk about learning: A national survey about young children and Science.* New York, NY, & Menlo Park, CA: Education Development Center, & SRI International.

Slavin, R. E., Lake, C., Hanley, P., & Thurston, A. (2014). Experimental evaluations of elementary science programs: A best-evidence synthesis. *Journal of Research in Science Teaching, 51*(7), 870–901. Doi: 10.1002/tea.21139

Strouse, G. A., O'Doherty, K., & Troseth, G. L. (2013). Effective coviewing: Preschoolers' learning from video after a dialogic questioning intervention. *Developmental Psychology, 49*(12), 2368.

Sweetman, S. S., Mirkin, L. S., Lund, A. E., & Bishop, S. K. (2018). Preschoolers learn to think and act like scientists with The *Cat in the Hat*. In S. Pasnik (Ed.), *Getting ready to learn: Creating effective, educational children's media* (pp. 118–134). New York, NY: Routledge.

Travis, R. (2017). What Homeric epic can teach us about educational affordances of interactive narrative. In M. F. Young & S. T. Slota (Eds.), *Exploding the castle: Rethinking how video games and game mechanics can shape the future of education* (pp. 19–37). Charlotte, NC: Information Age.

Troseth, G. L., Saylor, M. M., & Archer, A H.. (2006). Young children's use of video as a source of socially relevant information. *Child Development, 77*(3), 786–799.

Troseth, G. L. (2010). Is it life or is it Memorex? Video as a representation of reality. *Developmental Review, 30*(2), 155–175.

U.S. Department of Health and Human Services, Administration for Children and Families, Office of Head Start. (2015). *Head start early learning outcomes framework: Ages birth to five*. Retrieved from https://eclkc.ohs.acf.hhs.gov/sites/default/files/pdf/elof-ohs-framework.pdf

Woolley, J. D., & Ghossainy, M. E. (2013). Revisiting the fantasy-reality distinction: Children as naïve skeptics. *Child Development, 84*(5), 1496–1510.

Wright, T. S., & Neuman, S. B. (2014). Paucity and disparity in kindergarten oral vocabulary instruction. *Journal of Literacy Research, 46*(3), 330–357.

Young, M. F., Slota, S., Cutter, A. B., Jalette, G., Mullin, G., Lai, B., … Yukhymenko, M. (2012). Our princess is in another castle: A review of trends in serious gaming for education. *Review of Educational Research, 82*(1), 61–89.

Appendix A

Recruitment & Data Collection Procedures

Recruitment Procedures

The study team worked with Drive Research, a recruitment firm, to identify participants in five locations across the country: Boston, Massachusetts; Minneapolis, Minnesota; New York, New York; Phoenix, Arizona; and San Francisco, California. Recruiters recommended an initial list of approximately 10 locations based on the likelihood of finding large numbers of families within a 20-mile radius that fell below the local low-income threshold for a family of four according to the Section 8 guidelines established by the U.S. Department of Housing and Urban Development (HUD). The study team selected the final locations to provide regional diversity for the sample (Table A1). This sample is not meant to be representative of low-income children across the United States.

Table A1. Study Locations and the 2018 HUD Low-Income Limit Used for Classifying Families as "Low Income"

Market	Household Low-Income Limit
Boston, MA	$ 81,100
Minneapolis, MN	$ 71,900
New York, NY	$ 83,450
Phoenix, AZ	$ 55,300
San Francisco, CA	$ 117,400

Staff from Drive Research identified parents and caregivers (hereafter referred to as parents) through existing email lists and through Facebook advertisements. The Facebook advertisements briefly described the study, the research organizations involved, and incentives. Respondents interested in participating were invited to take a short online survey to ensure that they lived within the target area and met age and income requirements. Drive staff then made follow-up phone calls to qualified respondents to gather contact information, to provide additional information about the study and confirm interest, and to confirm that potential participants met the following criteria:

- Child was born between September 1, 2013 and September 30, 2014.
- Child is fluent in English.

- Household income is at or below 85% of the HUD low-income threshold for a family of four in theirareaasfoundathttps://www.huduser.gov/portal/datasets/il/il2018/select_Geography.odn.
- At least one parent is proficient in English or Spanish.
- Parent indicated that child could participate in game-like activities for up to 35 minutes (the duration of the pretest).

In response to challenges in identifying a sufficient number of eligible children, the study team decided to expand the age eligibility requirements from a 13-month to a 14-month range. To be eligible for the study, a child must have been born between September 1, 2013 – October 31, 2014. September 1, 2013 was chosen as final cutoff date because it is the kindergarten entry date for most of the participating locations.

Table A2. **Steps of the Recruitment Process**

Step 1	Step 2	Step 3	Step 4	Step 5
Outreach through Facebook and email lists to the parent(s) and guardian(s)	Participate in the prescreening online survey	Submit contact information	Rescreening phone call to confirm information for those qualified	Schedule participant family

The steps of the recruitment process are laid out above in Table A2. The recruitment team provided study information and conducted screening phone calls in English and Spanish. Families who did not have a child within the age range, who had more than one child within the age range, whose child did not speak and understand English, or who did not have a child who could reasonably participate in game-like activities for up to 35 minutes were screened out. If the family met the criteria and agreed to participate, they progressed to scheduling their meeting with the study team. To encourage retention and reduce attrition, the recruitment team also scheduled end-of-study meetings at the time of scheduling the first meeting. A total of 454 families gave written consent to participate in the study and successfully completed a baseline survey and child assessments.

Families attended one welcome meeting and one end-of-study meeting eight weeks apart; data collection took place over an approximately one-hour session at a prescheduled time. The meetings took place at a combination of Education Development Center (EDC) offices, Focus Pointe Global focus group facilities, and rented office spaces. In most locations, data collection took place over four days (Friday–Monday), so that scheduling was convenient for families.

Data Collection Procedures

Technology Setup Procedures

The study team provisioned data-enabled tablet computers in a random sequence as control or treatment tablets. The study team installed settings on all tablets to prevent children from accessing most age-inappropriate content through the Google Play Store (most restrictive settings for all types of media: Apps & Games, Movies, TV, Books, and Music).

Tablets for the control group featured the Chrome app and PDF versions of the Control Group Parent Guide in English and Spanish on the home screen. A set of apps and websites—including all PBS KIDS websites, science-related PBS KIDS apps, and *The Cat in the Hat Builds That*—were blocked for the duration of the study by using Screen Time security software. Parents in the control group were informed that some media would be blocked and that they would gain access after the second meeting. However, families may have been able to access *The Cat in the Hat Knows A Lot About That!* (*Cat in the Hat*) and other PBS KIDS shows through third-party websites, such as YouTube. The following apps were blocked for the control group:

- *The Cat in the Hat Builds That!*
- *Dinosaur Train Jurassic JR.*
- *Measure Up!*
- *PBS Parents Play and Learn Science*
- *PEEP Trash Stash*
- *PEEP Chirp Shapes Up*
- *PEEP Quack's Apples*
- *PEEP Family Science Ramps*
- *PEEP Family Science Rampas*
- *PBS KIDS Games*
- *PBS Kids Video*

Tablets for the treatment group had *The Cat in the Hat Builds That* and a custom-built *Cat in the Hat* video player bookmarked on the home screen. PDF versions of the Treatment Group Parent Guide in English and Spanish were also featured on the home screen. No apps or websites were blocked for the treatment group tablets.

Meeting 1

At the first meeting (Table A3), a member of the study team informed parents of study procedures and obtained their written consent to participate. Children then completed the Lens on Science, a computer-based assessment of children's basic science knowledge,[1] while parents completed a survey with information on child and family characteristics and their child's current use of digital media. If the child was unable or unwilling to complete the Lens assessment, the parent received a $25 gift card but the child was not enrolled in the study. This occurred nine times during the study.

[1] See the Data section in the main report for further description of the assessment.

The study team directed children who successfully completed the Lens assessment and their parents to a separate section of the facility, where they provided the parents with tablet computers. Before this meeting, members of the study team had randomly assigned each sequentially numbered tablet to either the treatment or control condition. Participants were randomly assigned to the treatment or control group through the sequential distribution of the numbered tablets. To avoid any bias on the part of the study team in the assessment room or of the study team member distributing tablets, Lens assessors affixed a number to each parent as the family completed the assessment step so that the tablet distributor could match the parent's number to the next tablet in sequence. The study organized the data collection procedures so that Lens assessors could not learn the treatment status of the tablet sequence before or after tablet distribution.

Table A3. **Steps at Meeting 1 for Treatment and Control Group Families**

	Step 1	Step 2	Step 3	Step 4
Treatment group	Consent conversation with parent and child assent	Lens assessment and parent survey	Tablet distribution (random assignment)	Orientation for treatment group
Control group				Orientation for control group

Each family received a canvas bag that included the tablet in a protective case, a power cord, an orientation packet, and a red or blue pen, which confirmed treatment or control group status to the study team. A study team member called families by name and brought them to a separate room for study orientation. One study team member provided orientation to the study for control-assigned families, and another study team member provided orientation for treatment-assigned families. Orientation sessions took place in separate corners of the facility to minimize interactions between treatment- and control-assigned families before and after orientation.

During the orientation session for the treatment group, the study team member encouraged participants to engage with the *Cat in the Hat* digital media content for one hour per week for eight weeks, or in keeping with their typical approach in guiding how much media the child used in a given day. The study team member also requested that only the child selected to participate in the study use the tablet for the duration of the study. Parents were told they would receive weekly text messages encouraging them to engage in the *Cat in the Hat* digital media content and real-world activities, reminding them to engage with the target theme for the week, and linking them to complete the weekly media log. Participants who did not fill out their weekly media log received another text reminder to do so. The treatment group's orientation packet included the following:

» A parent guide with information about the study, science-focused tips for parents that are also available on the *Cat in the Hat* website, and more about the *Cat in the Hat* digital media content. The guide also included troubleshooting tips for using the tablet.

» Three different *Cat in the Hat* real-world activities for parents and children to do together around the themes of the game and video content.

» A calendar with one of three randomized sequences of *Cat in the Hat* content to focus on each week (Figure A1). Each content theme had its own week, and there were two weeks of free choice.

During the orientation session for the control group, a study team member encouraged participants to engage with educational digital media content on the tablet for one hour per week for eight weeks, or in keeping with their typical approach in guiding how much media the child used in a given day. Their orientation packet included a parent guide with general information about the study, along with troubleshooting tips for using the tablet. Control group participants received no text messages during the study outside of those reminding them that the tablet would be remotely locked ahead of the second meeting.

At the end of each orientation session, families received $25 for participating in the first meeting. Families were encouraged to leave the facility at the conclusion of their orientation session to limit possible contamination across conditions.

Figure A1. Study Calendar Received by Treatment-Assigned Families

Meeting 2

On the Thursday preceding the second meeting (Table A4), the study team remotely locked all the study tablets, making it impossible for participants to use them. Parents were informed during Meeting 1 that this would happen and received a text message reminding them that a member of the study team would unlock the tablet when they arrived at the meeting. Locking the tablets ensured that all families in a specific location had nearly equal time with the materials, prevented families from revealing their condition to the study team members who were blind to condition by playing with the tablet onsite at Meeting 2, and provided an additional incentive to return for the second meeting.

Table A4. **Steps at Meeting 2 for Treatment and Control Group Families**

<table>
<tr><th></th><th>Step 1</th><th>Step 2</th><th>Step 3</th></tr>
<tr><td>Treatment group</td><td rowspan="2">Greeting and reacquaintance
Reconfirm child assent
Tablet collection for maintenance</td><td rowspan="2">Lens assessment and parent survey</td><td rowspan="2">Performance-based assessments</td></tr>
<tr><td>Control group</td></tr>
</table>

At Meeting 2, the study team greeted families and collected their tablets to unlock them, remove the tracking software, and delete proprietary PBS content. Children completed the Lens-Modified assessment, and parents filled out a survey asking about their child's learning activities over the last eight weeks. After completing the Lens-Modified assessment, children met with a member of the study team to complete two of the Hands-On Preschool Assessments of Physical Science and Engineering. (See Appendix D for further description of each of these assessments.) To limit the total assessment time to no more than 35 minutes, each child was randomly assigned two of the three assessment tasks. The specific performance tasks and the order in which the tasks were delivered were determined randomly based on the tablet ID before the initiation of the study. Assessors were provided with a ready-made assessment packet for each child that clearly designated the two tasks. Counterbalancing the order of assessment administration resulted in six different combinations of assessments to which a child could be assigned.

After assessments were complete, treatment- and control-assigned families received $100 gift cards and picked up their cleaned tablets. The tablets of control-assigned families who did not opt into the second study[2] were also returned to factory settings.

[2] At the time of consent, parents also were given the option to participate in a second study. Activities associated with the second study took place after the completion of this study. Families could participate in the first study whether or not they chose to participate in the second. The second study will be described in a separate report.

Appendix B

Analysis Details

Qualitative Analysis Details

The study team analyzed the parent interview data to identify common themes about how 4- to 5-year-old children and their parents engaged with *Cat in the Hat* materials. Before reviewing transcripts, the study team established four key categories: use/engagement (how the resources were used, parent/child perceptions of resources), challenges experienced, perceived benefits for parents, and perceived benefits for children. Two members of the study team first independently read each transcript and identified parent comments that related to each of the established categories. Coders next examined parent comments within each category, identifying subthemes that emerged and categorizing quotes into the subthemes. The coders then reviewed each other's subthemes, discussed any discrepancies, and came to a consensus on the key subthemes belonging to each category. The coders discussed which quotes were illustrative of the majority of parent comments for each of the four categories and selected key quotes to illustrate each of the subthemes. The Results section of the main report includes quotes as applicable to support quantitative results.

Quantitative Analysis Details

The following tables provide detail additional to what was included in the report, including sample demographic details, analytic sample characteristics, correlation tables, and regression tables.

Sample Characteristics

Table B1. Demographic and Baseline Characteristics of the Sample Overall and by Condition

	Total Sample (*N*=454)		Treatment Group (*N*=229)		Control Group (*N*=225)	
	N	% or mean *(SD)*	*N*	% or mean *(SD)*	*N*	% or mean *(SD)*
Site						
New York	106	23.3%				
Phoenix	116	25.6%				
Minneapolis	121	26.7%				
San Francisco	82	18.1%				
Boston	29	6.4%				
Lens pretest	449	1.37 (1.09)	227	1.39 (1.13)	222	1.34 (1.05)
Child is female	247	54.4%	128	55.9%	119	52.9%
Child age in months	454	59.2 (3.8)	229	59.1 (3.8)	225	59.2 (3.9)
Family income						
Less than $25,000	62	13.7%	28	12.2%	34	15.1%
$25,001 to $50,000	170	37.4%	83	36.2%	87	38.7%
$50,001 to $75,000	144	31.7%	73	31.9%	71	31.6%
$75,001 to $100,000	47	10.4%	29	12.7%	18	8.0%
$100,000 or more	20	4.4%	9	3.9%	11	4.9%
No response	11	2.4%	7	3.1%	4	1.8%
Parent education						
Less than high school diploma	8	1.8%	5	2.2%	3	1.3%
High school diploma or GED	70	15.4%	38	16.6%	32	14.2%
Some college (no degree)	141	31.1%	68	29.7%	73	32.4%

	Total Sample (*N*=454)		Treatment Group (*N*=229)		Control Group (*N*=225)	
	N	% or mean (*SD*)	*N*	% or mean (*SD*)	*N*	% or mean (*SD*)
Associate's or technical degree (AA, AS)	71	15.6%	34	14.8%	37	16.4%
Bachelor's degree (BA, BS)	115	25.3%	59	25.8%	56	24.9%
Graduate or professional degree	49	10.8%	25	10.9%	24	10.7%
Child race or ethnicity						
White	157	34.6%	75	32.8%	82	36.4%
Hispanic	85	18.7%	43	18.8%	42	18.7%
Black or African American	82	18.1%	42	18.3%	40	17.8%
Asian	15	3.3%	9	3.9%	6	2.7%
Other	11	2.4%	7	3.1%	4	1.8%
Multirace	104	22.9%	53	23.1%	51	22.7%
Relationship to child						
Parent	445	98.0%	224	97.8%	221	98.2%
Grandparent	6	1.3%	3	1.3%	3	1.3%
Other	3	0.7%	2	0.9%	1	0.4%
Parent/Caregiver responding to baseline survey is female	395	87.0%	194	84.7%	201	89.3%
Childcare or schooling						
Center-based preK or daycare	289	63.7%	149	66.2%	140	61.1%
Family or relative care	80	17.6%	36	16.0%	44	19.2%
Combination of center and family care	38	8.4%	20	8.9%	18	7.9%
Kindergarten	46	10.1%	19	8.4%	27	11.8%
No response	1	0.2%	1	0.4%	0	0.0%
Hours spent in center- or school-based preK, daycare, or K						
< 5 hours	7	1.5%	5	2.2%	2	0.9%
6 to 15 hours	107	23.6%	52	22.7%	55	24.4%

	Total Sample (*N*=454)		Treatment Group (*N*=229)		Control Group (*N*=225)	
	N	**% or mean (*SD*)**	***N***	**% or mean (*SD*)**	***N***	**% or mean (*SD*)**
16 to 30 hours	106	23.3%	54	23.6%	52	23.1%
More than 30 hours	153	33.7%	74	32.3%	79	35.1%
Not applicable (family care only)	81	17.8%	44	19.2%	37	16.4%
Care type & hours combo						
Any center care (non-K) 30+ hours	115	25.3%	64	28.4%	51	22.3%
Any center care (non-K) <30 hours	212	46.7%	105	46.7%	107	46.7%
Kindergarten, any hours	46	10.1%	19	8.4%	27	11.8%
Family/relative care only	80	17.6%	36	16.0%	44	19.2%
No response	1	0.2%	1	0.4%	0	0.0%
Languages spoken at home						
English only	342	75.3%	168	74.7%	174	76.0%
English and Spanish	73	16.1%	37	16.4%	36	15.7%
English and another language	25	5.5%	13	5.8%	12	5.2%
English and multiple languages	14	3.1%	7	3.1%	7	3.1%
Child's primary language at home						
Mostly English	90	19.8%	44	19.6%	46	20.1%
Mostly another language	5	1.1%	5	2.2%	0	0.0%
English and another language equally	17	3.7%	8	3.6%	9	3.9%
Not applicable	342	75.3%	168	74.7%	174	76.0%
Child has IEP	46	10.1%	19	8.3%	27	12.0%
No response	4	0.9%	3	1.3%	1	0.4%
Child has used *Cat in the Hat*	376	82.8%	184	81.8%	192	83.8%

Table B2. Child Baseline Science Knowledge and Demographic Characteristics, Overall and by Condition-Analytic Sample (Children Who Completed at least One Posttest Assessment)[3]

	Total Sample (*N*=443)	Treatment Group (*N*=223)	Control Group (*N*=220)
	% or mean ***(SD)***	**% or mean** ***(SD)***	**% or mean** ***(SD)***
Child is female	54.9%	56.5%	53.2%
Child age in months	59.2 (3.9)	59.1 (3.9)	59.3 (3.9)
Baseline science knowledge (*Lens on Science*)	1.4 (1.1)	1.4 (1.1)	1.4 (1.1)
Child race or ethnicity			
White	34.5%	32.3%	36.8%
Hispanic	18.3%	18.8%	17.7%
Black or African American	18.3%	18.8%	17.7%
Asian	3.4%	4.0%	2.7%
Other	2.5%	3.1%	1.8%
Multirace	23.0%	22.9%	23.2%
Child attends center care (non-K) for 30+ hours	25.3%	22.4%	28.2%
Child's primary language is English	75.9%	76.2%	75.5%
Child has an IEP or 504 plan	9.9%	8.1%	11.8%
Family income more than $50K	47.2%	49.3%	45.0%
Responding parent has high school education or lower	82.6%	80.7%	84.6%

[3] Table B2 displays information on only those demographic characteristics that were used as covariates in the analytic models. Table B1 displays information on all study collected demographic characteristics.

Table B3. Baseline Balance between Treatment and Control Groups on Lens Pretest

	Overall		Treatment Group		Control Group		
	N	Mean (*SD*)	*N*	Mean (*SD*)	*N*	Mean (*SD*)	Effect size (*g*)
Lens pretest	449	1.37 (1.09)	227	1.39 (1.13)	222	1.34 (1.05)	-0.048

Correlation Table

Table B4. Bivariate Correlations between All Outcome Measures and Selected Child and Family Demographic Characteristics

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
		Lens pre	Lens post	Length, Strength, & Stability	Surfaces & Friction	Child female	Child age	Family inc. $50K+	Parent ed. HS diploma+	Child is white	FT center care	K.	Eng. only at hm.	Has IEP	Prior *Cat in the Hat*
1	Lens pretest	--				.08~	.38***	.09~	.14**	.23***	-.08~	.20***	.14**	-.08	.07
2	Lens posttest	.68***	--			.05	.34***	.08	.09~	.22***	-.12*	.16***	.11*	-.12*	.06
3	Length, Strength, & Stability	.41***	.40***	--		.02	.15*	.06	.13*	.13*	.05	-.04	.12*	-.10~	-.02
4	Surfaces & Friction	.50***	.50***	.45***	--	.14*	.21***	.01	.11~	.16**	-.13*	.09	.13*	-.01	.05
5	Colors, Shapes, & Uses	.44***	.51***	.37***	.44***	.06	.22***	.10	.11~	.11~	-.02	.10	.10	-.03	-.00

Note. inc. = income, HS = high school, FT = full-time, K = kindergarten, Eng. = English. “Full-time center care” refers to any center-based preK, non-kindergarten attended for at least 30 hours per week.

~$p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Regression Models

Table B5. Regression Models Predicting Lens-Modified Posttest Scores with Treatment Status, Lens Pretest, and Demographic Covariates

	Model 1				Model 2				Model 3			
	b	*(SE)*	*p*	*d*	*b*	*(SE)*	*p*	*d*	*b*	*(SE)*	*p*	*d*
Cat in the Hat (treatment)	0.18	(0.10)	0.07	0.17	0.11	(0.07)	0.12	0.11	0.11	(0.07)	0.14	0.11
Lens pretest					0.65	(0.03)	<0.001		0.57	(0.04)	<0.001	
Child age in months									0.04	(0.01)	<0.001	
Parent education, HS or higher									0.02	(0.10)	0.87	
Child attends center-based care, 30+ hours									-0.18	(0.09)	0.04	
Child attends kindergarten									-0.02	(0.13)	0.89	
Child has IEP									-0.35	(0.12)	0.01	
Child is white									0.17	(0.08)	0.05	
English only at home									0.03	(0.09)	0.74	
Intercept	1.52	(0.07)			0.63	(0.07)			-1.37	(0.64)	0.03	
R-squared	0.01				0.47				0.50			
Adjusted *R*-squared	0.01				0.46				0.49			
N	432				427				426			

Note. d = Cohen's *d* effect size.

Table B6. **Regression Models Predicting Length, Strength, and Stability Assessment Scores with Treatment Status, Lens Pretest, and Demographic Covariates**

	Model 1				Model 2				Model 3			
	b	*(SE)*	*p*	*d*	*b*	*(SE)*	*p*	*d*	*b*	*(SE)*	*p*	*d*
Cat in the Hat (treatment)	0.84	(0.23)	<0.001	0.43	0.79	(0.21)	<0.001	0.40	0.83	(0.21)	<0.001	0.42
Lens pretest					0.73	(0.10)	<0.001		0.63	(0.11)	<0.001	
Child age in months									0.02	(0.03)	0.47	
Parental education, HS diploma+									0.42	(0.28)	0.12	
Child is white									0.19	(0.24)	0.43	
English only at home									0.15	(0.25)	0.54	
Has IEP									-0.42	(0.36)	0.22	
Intercept	3.86	(0.16)			2.87	(0.20)	<0.001		1.52	(1.77)	0.39	
R-squared	0.05				0.21				0.22			
Adjusted *R*-squared	0.04				0.20				0.20			
N	287				283				283			

Note. *d* = Cohen's *d* effect size.

Table B7. **Regression Models Predicting Surfaces and Friction Assessment Scores with Treatment Status, Lens Pretest, and Demographic Covariates**

	Model 1				Model 2				Model 3			
	b	*(SE)*	*p*	*d*	*b*	*(SE)*	*p*	*d*	*b*	*(SE)*	*p*	*d*
Cat in the Hat (treatment)	1.13	(0.40)	0.004	0.34	1.11	(0.34)	0.001	0.33	1.11	(0.35)	0.002	0.33
Lens pretest					1.49	(0.15)	<0.001		1.35	(0.18)	<0.001	
Child gender									0.75	(0.35)	0.03	
Child age in months									0.03	(0.05)	0.59	
Parental education, HS diploma+									0.43	(0.46)	0.35	
Child is white									0.19	(0.39)	0.63	
Child attends center-based preK 30+ hours									-0.46	(0.41)	0.27	
English only at home									0.31	(0.42)	0.47	
Child has IEP									0.07	(0.60)	0.91	
Intercept	5.49	(0.28)	<0.001		3.36	(0.32)	<0.001		0.94	(3.11)	0.76	
R-squared	0.03				0.28				0.30			
Adjusted *R*-squared	0.03				0.28				0.28			
N	286				282				282			

Note. *d* = Cohen's *d* effect size.

Table B8. **Regression Models Predicting Performance-Based Assessment Colors, Shapes, and Uses Scores with Treatment Status, Lens Pretest, and Demographic Covariates**

	Model 1				Model 2				Model 3			
	b	*(SE)*	*p*	*d*	*b*	*(SE)*	*p*	*d*	*b*	*(SE)*	*p*	*d*
Cat in the Hat (treatment)	0.03	(0.03)	0.24	0.14	0.03	(0.02)	0.18	0.15	0.03	(0.02)	0.17	0.15
Lens pretest					0.09	(0.01)	<0.001		0.08	(0.01)	<0.001	
Demographics												
Child age in months.									0.01	(0.00)	0.15	
Parental education, HS diploma+									0.03	(0.03)	0.29	
Child is white									0.02	(0.02)	0.36	
Child has IEP									-0.00	(0.04)	0.90	
Intercept	0.53	(0.02)	<0.001		0.41	(0.02)	<0.001		0.12	(0.18)	0.51	
R-squared	0.01				0.20				0.21			
Adjusted *R*-squared	0.00				0.19				0.19			
N	275				273				273			

Note. *d* = Cohen's *d* effect size.

Interaction Regression Models

Table B9. Regression Models Predicting Lens-Modified Posttest Interacting Treatment Status with Lens Pretest and Demographics

	Model 1		Model 2		Model 3		Model 4	
	Treatment x Lens pretest		Treatment x child age		Treatment x parent ed.		Treatment x child female	
	b (*SE*)	*p*	*b* (*SE*)	*p*	*b* (*SE*)	*p*	*b* (*SE*)	*p*
Cat in the Hat **(treatment)**	0.11 (0.07)	0.12	0.12 (0.07)	0.10	0.24 (0.18)	0.18	0.07 (0.11)	0.54
Lens pretest	0.65 (0.05)	<0.001	0.61 (0.04)	<0.001	0.65 (0.03)	<0.001	0.65 (0.03)	<0.001
Child age in months			0.03 (0.01)	0.04				
Parent education, HS diploma+					0.10 (0.14)	0.50		
Child is female							-0.03 (0.10)	0.78
Treatment x Lens pre	-0.00 (0.07)	1.00						
Treatment x child age in months			0.01 (0.02)	0.66				
Treatment x parent ed., HS+					-0.15 (0.19)	0.43		
Treatment x child female							0.08 (0.15)	0.57
Intercept	1.52 (0.05)	<0.001	1.51 (0.05)	<0.001	1.43 (0.13)	<0.001	1.53 (0.08)	<0.001
***R*-squared**	0.47		0.48		0.47		0.47	
Adjusted *R*-squared	0.46		0.47		0.46		0.46	
N	427		427		427		427	

Note. Lens pretest and child age in months are grand mean centered.

Table B9. Regression Models Predicting Lens Posttest Interacting Treatment with Lens Pretest and Demographics (continued)

	Model 5		Model 6		Model 7		Model 8	
	Treatment x child white		Treatment x Eng. only		Treatment x center care 30+ hrs		Treatment x kindergarten	
	b (*SE*)	*p*	*b* (*SE*)	*p*	*b* (*SE*)	*p*	*b* (*SE*)	*p*
Cat in the Hat (treatment)	0.13 (0.09)	0.14	-0.02 (0.15)	0.89	0.03 (0.08)	0.69	0.12 (0.08)	0.12
Lens pretest	0.63 (0.04)	<0.001	0.65 (0.03)	<0.001	0.64 (0.03)	<0.001	0.64 (0.04)	<0.001
Child is white	0.17 (0.11)	0.13						
Child speaks English only at home			-0.05 (0.12)	0.71				
Child attends center-based care 30+ hrs					-0.32 (0.12)	0.01		
Child attends kindergarten							0.17 (0.18)	0.35
Treatment x child white	-0.02 (0.15)	0.89						
Treatment x English only			0.18 (0.17)	0.31				
Treatment x center-based care 30+ hrs					0.27 (0.17)	0.12		
Treatment x kindergarten							-0.13 (0.24)	0.60
Intercept	1.45 (0.07)	<0.001	1.55 (0.11)	<0.001	1.61 (0.06)	<0.001	1.50 (0.05)	<0.001
R-squared	0.47		0.47		0.48		0.47	
Adjusted *R*-squared	0.47		0.46		0.47		0.46	
N	427		427		426		426	

Note. Lens pretest and child age in months are grand mean centered.

Table B10. **Regression Models Predicting Length, Strength, and Stability Assessment Scores Interacting Treatment Status with Lens Pretest and Demographics**

	Model 1		Model 2		Model 3		Model 4	
	Treatment x Lens pretest		Treatment x child age		Treatment x parent ed.		Treatment x child female	
	b (*SE*)	*p*	*b* (*SE*)	*p*	*b* (*SE*)	*p*	*b* (*SE*)	*p*
Cat in the Hat **(treatment)**	0.78 (0.21)	<0.001	0.79 (0.21)	<0.001	0.52 (0.50)	0.30	0.98 (0.31)	<0.001
Lens pretest	0.61 (0.14)	<0.001	0.71 (0.11)	<0.001	0.69 (0.10)	<0.001	0.73 (0.10)	<0.001
Child age in months			0.02 (0.04)	0.53				
Parent education, HS diploma+					0.21 (0.42)	0.62		
Child is female							0.14 (0.29)	0.63
Treatment x Lens pretest	0.23 (0.20)	0.25						
Treatment x child age in months			-0.03 (0.05)	0.61				
Treatment x parent ed., HS+					0.37 (0.55)	0.50		
Treatment x child female							-0.37 (0.42)	0.38
Intercept	-0.43 (0.15)	<0.001	-0.42 (0.15)	0.01	-0.60 (0.39)	0.13	-0.49 (0.21)	0.02
***R*-squared**	0.21		0.21		0.21		0.21	
Adjusted *R*-squared	0.20		0.20		0.20		0.20	
N	283		283		283		283	

Note. Lens pretest and child age in months are grand mean centered.

Table B10. **Regression Models Predicting Length, Strength, and Stability Assessment Scores Interacting Treatment Status with Lens Pretest and Demographics (continued)**

	Model 5		Model 6		Model 7		Model 8	
	Treatment x child white		Treatment x Eng. only		Treatment x center care 30+ hrs		Treatment x kindergarten	
	b *(SE)*	*p*	*b* *(SE)*	*p*	*b* *(SE)*	*p*	*b* *(SE)*	*p*
Cat in the Hat (treatment)	0.863 (0.26)	<0.001	0.55 (0.41)	0.18	0.67 (0.24)	0.01	0.90 (0.22)	<0.001
Lens pretest	0.71 (0.10)	<0.001	0.71 (0.10)	<0.001	0.74 (0.10)	<0.001	0.76 (0.10)	<0.001
Child is white	0.28 (0.32)	0.37						
Child speaks English only at home			0.04 (0.34)	0.91				
Child attends center-based care 30+ hours					0.19 (0.32)	0.56		
Child attends kindergarten							-0.24 (0.51)	0.63
Treatment x child white	-0.21 (0.44)	0.63						
Treatment x English only			0.32 (0.48)	0.50				
Treatment x center-based care 30+ hours					0.51 (0.47)	0.28		
Treatment x kindergarten							-1.01 (0.67)	0.14
Intercept	-0.52 (0.18)	0.005	-0.45 (0.29)	0.13	-0.47 (0.18)	0.01	-0.39 (0.15)	0.01
R-squared	0.21		0.21		0.22		0.23	
Adjusted *R*-squared	0.20		0.20		0.21		0.22	
N	283		283		282		282	
Note. Lens pretest and child age in months are grand mean centered.								

Table B11. **Regression Models Predicting Surfaces and Friction Assessment Scores Interacting Treatment Status with Lens Pretest and Demographics**

	Model 1		Model 2		Model 3		Model 4	
	Treatment x Lens pre		Treatment x age		Treatment x parent ed.		Treatment x child female	
	b (*SE*)	*p*	*b* (*SE*)	*p*	*b* (*SE*)	*p*	*b* (*SE*)	*p*
Cat in the Hat (treatment)	1.10 (0.34)	<0.001	1.10 (0.34)	<0.001	0.68 (0.82)	0.41	1.13 (0.51)	0.03
Lens pretest	1.38 (0.22)	<0.001	1.45 (0.17)	<0.001	1.47 (0.15)	<0.001	1.47 (0.15)	<0.001
Child age in months			-0.03 (0.07)	0.66				
Parent education, HS diploma+					0.06 (0.65)	0.93		
Child is female							0.80 (0.49)	0.10
Treatment x Lens pretest	0.20 (0.30)	0.50						
Treatment x child age in months			0.09 (0.09)	0.30				
Treatment x parent ed., HS+					0.52 (0.91)	0.57		
Treatment x child female							-0.11 (0.69)	0.87
Intercept	-0.66 (0.24)	0.01	-0.65 (0.25)	0.01	-0.71 (0.59)	0.23	-1.08 (0.35)	<0.001
R-squared	0.28		0.28		0.28		0.29	
Adjusted *R*-squared	0.27		0.27		0.27		0.28	
N	282		282		282		282	

Note. Lens pretest and child age in months are grand mean centered.

Table B11. Regression Models Predicting Surfaces and Friction Assessment Scores Interacting Treatment Status with Lens Pretest and Demographics (continued)

	Model 5		Model 6		Model 7		Model 8	
	Treatment x child white		Treatment x Eng. only		Treatment x center care 30+ hrs		Treatment x kindergarten	
	b (*SE*)	*p*	*b* (*SE*)	*p*	*b* (*SE*)	*p*	*b* (*SE*)	*p*
Cat in the Hat (treatment)	1.17 (0.43)	0.01	1.17 (0.71)	0.10	1.06 (0.39)	0.01	1.21 (0.37)	<0.001
Lens pretest	1.46 (0.16)	<0.001	1.46 (0.15)	<0.001	1.46 (0.15)	<0.001	1.50 (0.15)	<0.001
Child is white	0.32 (0.52)	0.54						
Child speaks English only at home			0.45 (0.59)	0.45				
Child attends center-based care 30+ hours					-0.66 (0.57)	0.25		
Child attends kindergarten							0.35 (0.85)	0.68
Treatment x child white	-0.09 (0.73)	0.90						
Treatment x English only			-0.07 (0.81)	0.93				
Treatment x center-based care 30+ hours					0.19 (0.80)	0.81		
Treatment x kindergarten							-0.83 (1.10)	0.45
Intercept	-0.79 (0.32)	0.01	-1.01 (0.52)	0.05	-0.50 (0.28)	0.08	-0.70 (0.26)	0.01
R-squared	0.28		0.28		0.29		0.28	
Adjusted *R*-squared	0.27		0.27		0.28		0.27	
N	282		282		282		282	

Note. Lens pretest and child age in months are grand mean centered.

Table B12. Regression Models Predicting Colors, Shapes, and Uses Assessment Scores Interacting Treatment Status with Lens Pretest and Demographics

	Model 1		Model 2		Model 3		Model 4	
	Treatment x Lens pre		Treatment x child age		Treatment x parent ed.		Treatment x child female	
	b (*SE*)	*p*	*b* (*SE*)	*p*	*b* (*SE*)	*p*	*b* (*SE*)	*p*
***Cat in the Hat* (treatment)**	0.03 (0.02)	0.20	0.03 (0.02)	0.21	-0.03 (0.05)	0.65	0.06 (0.03)	0.07
Lens pretest	0.07 (0.02)	<0.001	0.08 (0.01)	<0.001	0.09 (0.01)	<0.001	0.09 (0.01)	<0.001
Child age in months			0.00 (0.00)	0.47				
Parent education, HS diploma+					-0.01 (0.05)	0.87		
Child is female							0.04 (0.03)	0.22
Treatment x Lens pretest	0.03 (0.02)	0.20						
Treatment x child age in months			0.00 (0.01)	0.69				
Treatment x parent ed., HS+					0.07 (0.06)	0.25		
Treatment x child female							-0.06 (0.05)	0.22
Intercept	-0.02 (0.02)	0.22	-0.02 (0.02)	0.23	-0.01 (0.04)	0.73	-0.04 (0.03)	0.08
***R*-squared**	0.20		0.20		0.20		0.20	
Adjusted *R*-squared	0.19		0.19		0.19		0.19	
N	273		273		273		273	

Note. Lens pretest and child age in months are grand mean centered

Table B12. Regression Models Predicting Colors, Shapes, and Uses Assessment Scores Interacting Treatment Status with Lens Pretest and Demographics (continued)

	Model 5		Model 6		Model 7		Model 8	
	Treatment x child white		Treatment x Eng. only		Treatment x center care 30+ hrs		Treatment x kindergarten	
	b (*SE*)	*p*	*b* (*SE*)	*p*	*b* (*SE*)	*p*	*b* (*SE*)	*p*
Cat in the Hat (treatment)	0.02 (0.03)	0.48	-0.10 (0.05)	0.04	-0.00 (0.03)	0.94	0.03 (0.02)	0.23
Lens pretest	0.08 (0.01)	<0.001	0.08 (0.01)	<0.001	0.09 (0.01)	<0.001	0.09 (0.01)	<0.001
Child is white	0.01 (0.03)	0.88						
Child speaks English only at home			-0.05 (0.04)	0.15				
Child attends center-based care 30+ hours					-0.05 (0.03)	0.12		
Child attends kindergarten							0.02 (0.06)	0.78
Treatment x child white	0.03 (0.05)	0.50						
Treatment x English only			0.16 (0.05)	0.00				
Treatment x center-based care 30+ hours					0.11 (0.05)	0.03		
Treatment x kindergarten							-0.01 (0.07)	0.87
Intercept	-0.02 (0.02)	0.28	0.02 (0.03)	0.51	-0.00 (0.02)	0.92	-0.02 (0.02)	0.23
R-squared	0.20		0.23		0.21		0.20	
Adjusted *R*-squared	0.19		0.21		0.20		0.18	
N	273		273		272		272	
Note. Lens pretest and child age in months are grand mean centered.								

Parent survey regression tables

Table B13. Regression Models Predicting Total Activities Engaged in Monthly or More with Treatment Status, Lens Pretest, and Demographics

	Model 1				Model 2				Model 3			
	b	*(SE)*	*d*	*p*	*b*	*(SE)*	*d*	*p*	*b*	*(SE)*	*d*	*p*
Cat in the Hat (treatment)	0.36	(0.17)	0.20	0.04	0.36	(0.17)	0.20	0.04	0.32	(0.17)	0.18	0.06
Lens pretest					0.04	(0.08)		0.59	0.03	(0.08)		0.68
Demographics												
Child is female									-0.19	(0.17)		0.28
Family income $50K+									0.26	(0.17)		0.14
Prior *Cat in the Hat*									0.65	(0.23)		<0.001
Intercept	9.41	(0.12)		<0.001	9.35	(0.16)		<0.001	8.80	(0.27)		<0.001
R-squared	0.01				0.01				0.03			
Adjusted *R*-squared	0.01				0.01				0.02			
N	444				439				428			

Note. d = Cohen's *d* effect size.

Table B14. **Regression Models Predicting Excitement about Reading with Treatment Status, Lens Pretest, and Demographics**

	Model 1				Model 2				Model 3			
	b	*(SE)*	*d*	*p*	*b*	*(SE)*	*d*	*p*	*b*	*(SE)*	*d*	*p*
Cat in the Hat **(treatment)**	-0.14	(0.18)	-0.07	0.44	-0.15	(0.18)	-0.08	0.42	-0.15	(0.18)	-0.08	0.41
Lens pretest					0.11	(0.08)		0.20	0.07	(0.09)		0.42
Demographics												
Child is female									0.49	(0.19)		0.01
Family income $50K+									0.33	(0.19)		0.07
Prior *Cat in the Hat*									0.10	(0.25)		0.67
Intercept	7.78	(0.13)		<0.001	7.63	(0.17)		<0.001	7.17	(0.29)		<0.001
***R*-squared**	0.00				0.01				0.03			
Adjusted *R*-squared	-0.00				0.00				0.02			
N	443				438				427			

Note. d = Cohen's *d* effect size.

Table B15. Regression Models Predicting Excitement about Science with Treatment Status, Lens Pretest, and Demographics

	Model 1				Model 2				Model 3			
	b	*(SE)*	*d*	*p*	*b*	*(SE)*	*d*	*p*	*b*	*(SE)*	*d*	*p*
***Cat in the Hat* (treatment)**	0.52	(0.20)	0.24	0.01	0.50	(0.20)	0.24	0.01	0.50	(0.20)	0.23	0.01
Lens pretest					0.06	(0.09)		0.55	0.07	(0.09)		0.46
Demographics												
Child is female									-1.09	(0.20)		<0.001
Family income $50K+									0.45	(0.20)		0.03
Prior *Cat in the Hat*									0.36	(0.27)		0.17
Intercept	6.91	(0.14)		<0.001	6.83	(0.19)		<0.001	6.88	(0.31)		<0.001
***R*-squared**	0.02				0.02				0.09			
Adjusted *R*-squared	0.01				0.01				0.08			
N	444				439				428			

Note. d = Cohen's *d* effect size.

Table B16. **Regression Models Predicting Excitement about Engineering with Treatment Status, Lens Pretest, and Demographics**

	Model 1				Model 2				Model 3			
	b	*(SE)*	*d*	*p*	*b*	*(SE)*	*d*	*p*	*b*	*(SE)*	*d*	*p*
Cat in the Hat **(treatment)**	0.07	(0.14)	0.04	0.65	0.08	(0.15)	0.05	0.61	0.07	(0.15)	0.04	0.66
Lens pretest					-0.04	(0.07)		0.53	-0.04	(0.07)		0.52
Demographics												
Child is female									-0.10	(0.15)		0.51
Family income $50K+									-0.15	(0.15)		0.33
Prior *Cat in the Hat*									0.32	(0.20)		0.11
Intercept	8.70	(0.10)		<0.001	8.76	(0.14)		<0.001	8.61	(0.23)		<0.001
***R*-squared**	0.00				0.00				0.01			
Adjusted *R*-squared	-0.00				-0.00				-0.00			
N	444				439				428			

Note. *d* = Cohen's *d* effect size.

Table B17. **Regression Models Predicting Study-Related Vocabulary Words with Treatment Status, Lens Pretest, and Demographics**

	Model 1				Model 2				Model 3			
	b	*(SE)*	*d*	*p*	*b*	*(SE)*	*d*	*p*	*b*	*(SE)*	*d*	*p*
Cat in the Hat **(treatment)**	0.59	(0.31)	0.18	0.06	0.58	(0.31)	0.18	0.06	0.45	(0.31)	0.14	0.15
Lens pretest					0.50	(0.14)		0.00	0.51	(0.14)		0.00
Demographics												
Child is female									-0.20	(0.31)		0.53
Family income $50K+									0.69	(0.31)		0.03
Prior ***Cat in the Hat***									-0.03	(0.42)		0.95
Intercept	8.58	-0.22		<0.001	7.87	-0.29		<0.001	7.69	(0.49)		<0.001
***R*-squared**	0.01				0.04				0.05			
Adjusted ***R*-squared**	0.01				0.03				0.04			
N	444				439				428			

Note. *d* = Cohen's *d* effect size.

Table B18. **Regression Models Predicting Placebo Vocabulary Words with Treatment Status, Lens Pretest, and Demographics**

	Model 1				Model 2				Model 3			
	b	*(SE)*	*d*	*p*	*b*	*(SE)*	*d*	*p*	*b*	*(SE)*	*d*	*p*
Cat in the Hat **(treatment)**	-0.10	(0.10)	-0.09	0.33	-0.10	(0.10)	-0.10	0.31	-0.11	(0.10)	-0.11	0.24
Lens pretest					0.22	(0.04)		0.00	0.22	(0.04)		0.00
Demographics												
Child is female									-0.19	(0.10)		0.05
Family income $50K+									0.07	(0.10)		0.49
Prior *Cat in the Hat*									0.10	(0.13)		0.43
Intercept	1.57	-0.07		<0.001	1.26	-0.09		<0.001	1.25	(0.15)		<0.001
***R*-squared**	0.00				0.05				0.07			
Adjusted *R*-squared	-0.00				0.05				0.06			
N	444				439				428			

Note. *d* = Cohen's *d* effect size.

Appendix C

Additional Study Materials

Real-World Activities

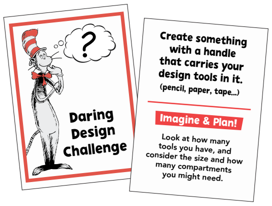
The following activities were provided to treatment-assigned families:

- Daring Design Challenge
- Measuring This and That
- What Floats Your Boat?

Daring Design Challenge

Work together to build something new.

Sally and Nick build a bridge that's long and strong, a boat that floats, and a vase with a cool base. What can YOU create? Work together with a friend and challenge your building brains by designing some amazing creations in this collaborative game.



Materials

- Daring Design Challenge cards (print two-sided)
- The Engineering Design Process wheel
- Paper
- Pencil
- Tape
- Various household items
  (Examples: wooden craft sticks, paper, straws, building blocks, fabric scraps, cans, aluminum foil, paper plates and cups, string, etc.)

Play the Game

1. Print and cut out the Daring Design Challenge cards and the Engineering Design Process wheel.
2. Together, select a game card from the pile.
3. Follow the challenge and make something amazing as you move step by step—along with The Cat in the Hat—through the Design Process wheel.

Tip

The shape and weight of a design can affect how it works. You may need a wider base or different material for your designs to balance, float, or move.

Remember

Accidents happen — if you're lucky! It's okay if things don't always happen how we planned. Sometimes we discover something unexpected that makes our designs even better.

More Ways to Play

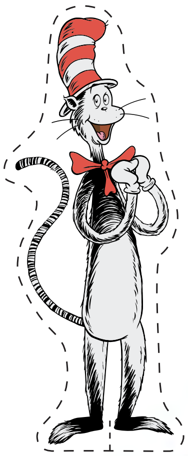
1. Think of a problem that needs to be solved. Using the four blank Daring Design Challenge cards (included in this printout), create four more design challenges. Try them with a friend.
2. Choose a design you already created and make it again using a different material.
3. Time how long it takes you to build a design. Then try it again. Can you beat your time?

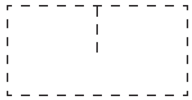
Find related games in the FREE *The Cat in the Hat Builds That* app. Download it now!

pbskids.org/catinthehat

The contents of this document were developed under a cooperative agreement (PR/Award No. U295A150003, CFDA No. 84.295A) from the U.S. Department of Education. However, these contents do not necessarily represent the policy of the Department of Education, and you should not assume endorsement by the Federal Government.

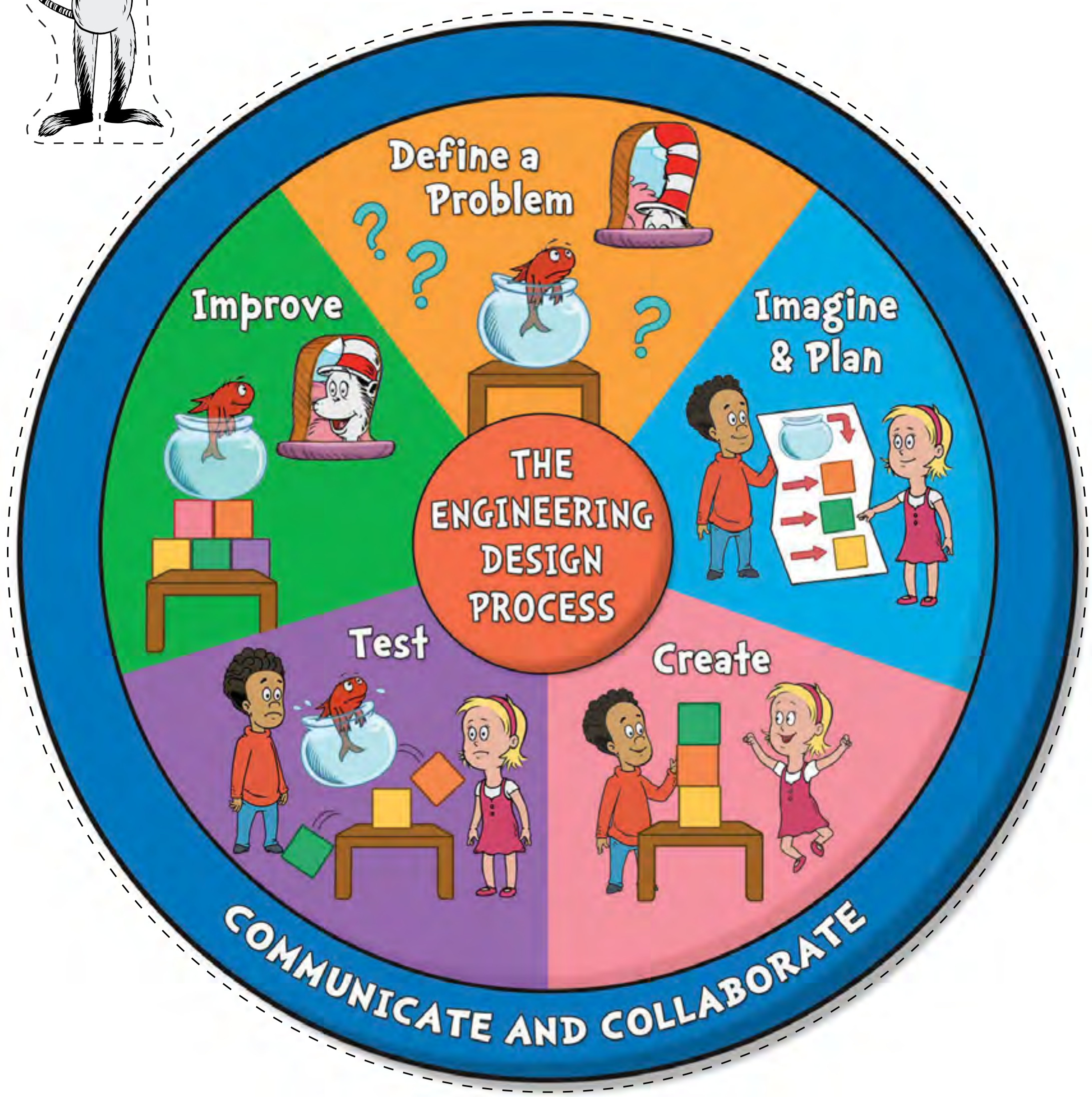
The Engineering Design Process



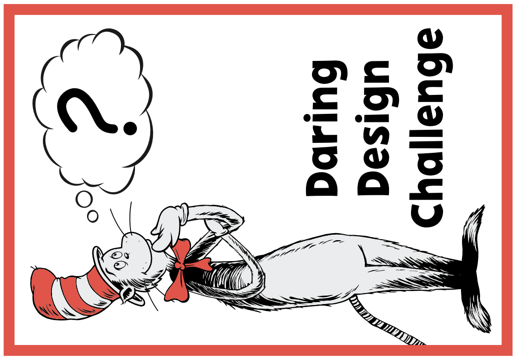


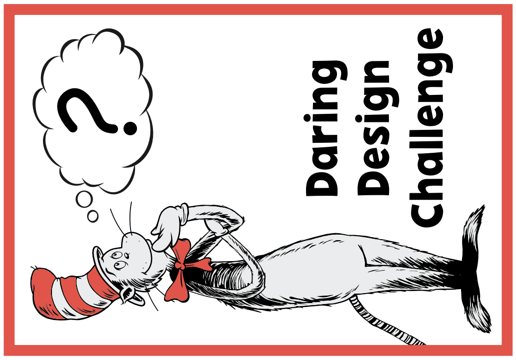
Stabilizer

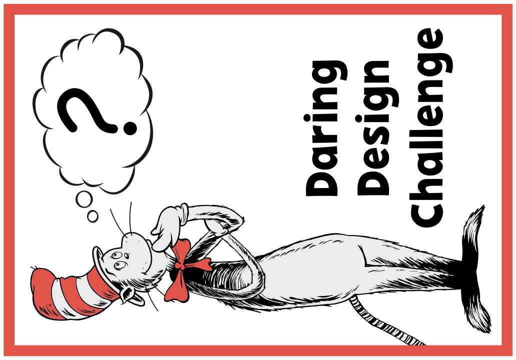
1. With an adult's help, cut out the Cat figure, the stabilizer, and the wheel.
   Then cut along the vertical lines at the base of the Cat and the top of the stabilizer.
2. Push the stabilizer into the base of the Cat to create a "+". The Cat should now stand.
3. Use the Cat to guide you around the wheel.

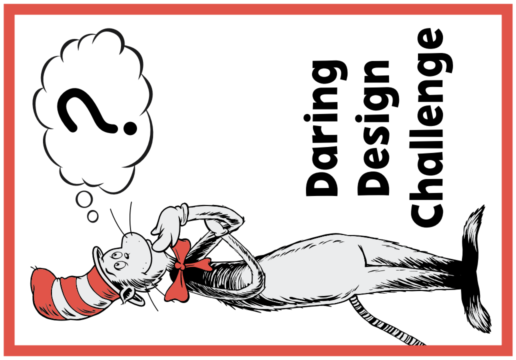


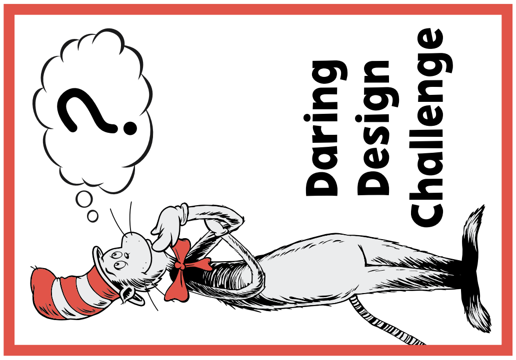
Print two-sided

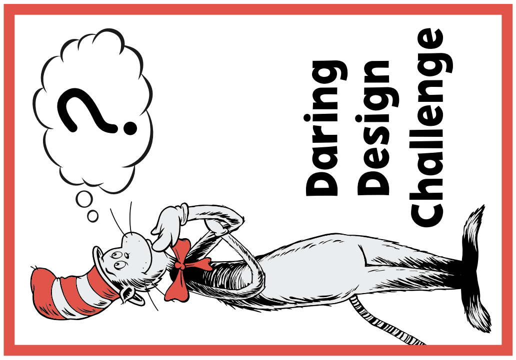


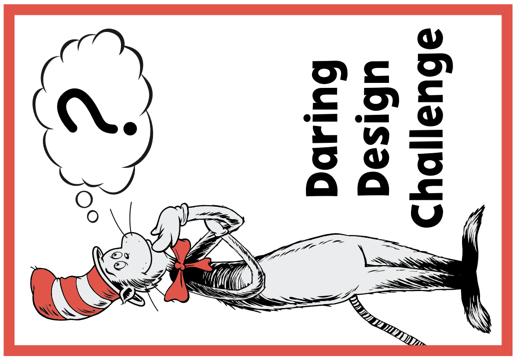


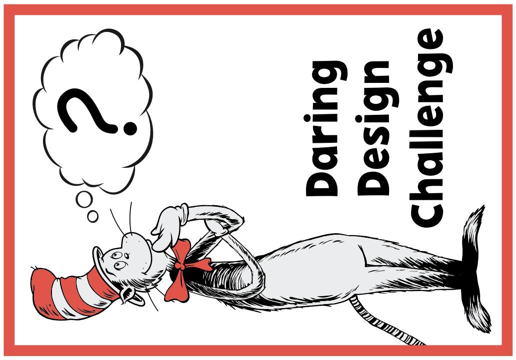












Make something that balances on one or two legs.

Improve!

What if you bumped your creation?
What happens?
Can you make it better?

Make something that balances on three legs.

Create!

Work together as you build and listen to all ideas.
This is called Collaboration—just like the three legs are collaborating to balance your design!

Design a hat that makes you taller than an adult (or the Cat in the Hat!).

Test!

What can you use to measure and get evidence?
Are you taller?

Make a house for Fish that doesn't float.

Imagine & Plan!

What materials and shapes DON'T float?
How can you use them in your design?

Design something that makes it easier to brush your teeth.

Define the Problem!

Toothbrush hard to hold?
Difficult getting the last bit of toothpaste out of the tube?
Or something else?

Make something that rolls the length of your height when pushed or pulled.

Improve!

Can you double that distance?

Design something to wear that keeps you warm.

Test!

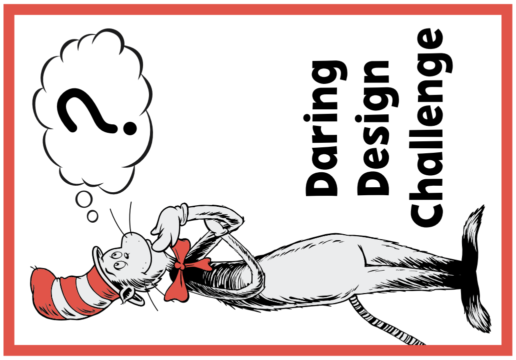
Try standing in front of your open freezer.
How do you feel?

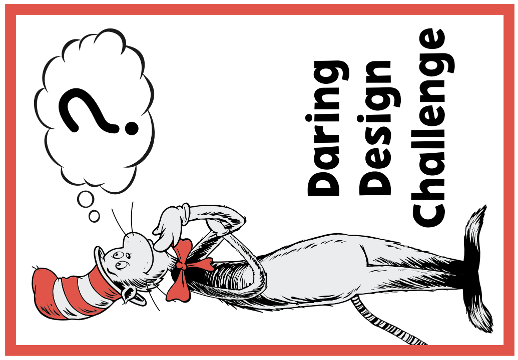
Make a bridge as long as your arm.

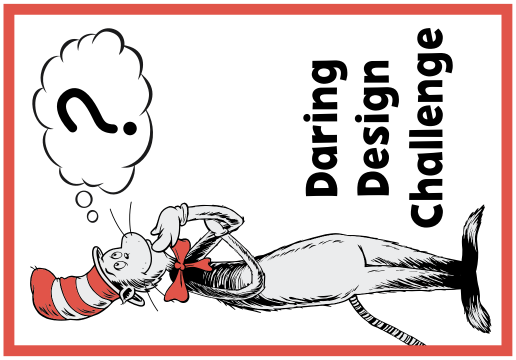
Test!

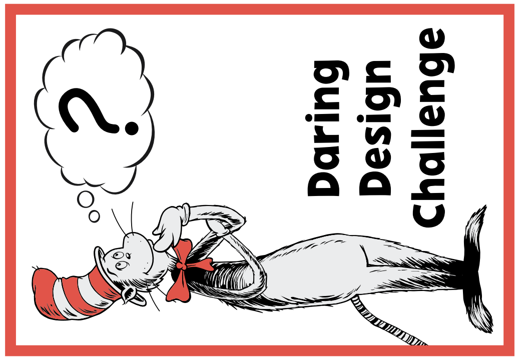
What if you put a penny or a spoon on it?
How much weight can it take?

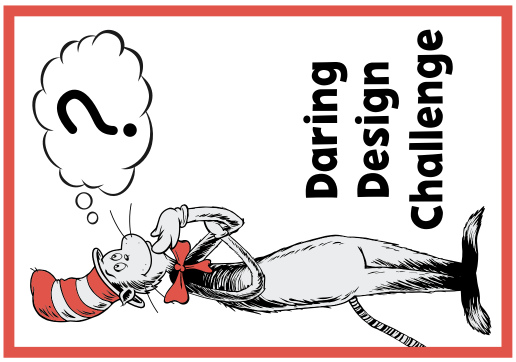
Print two-sided

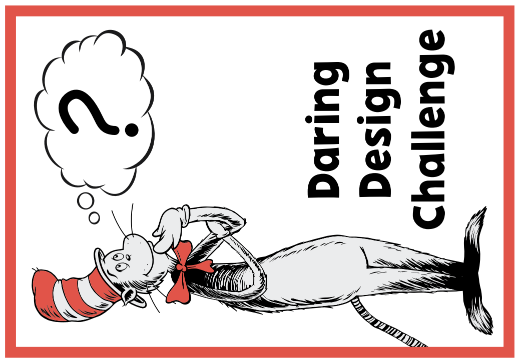


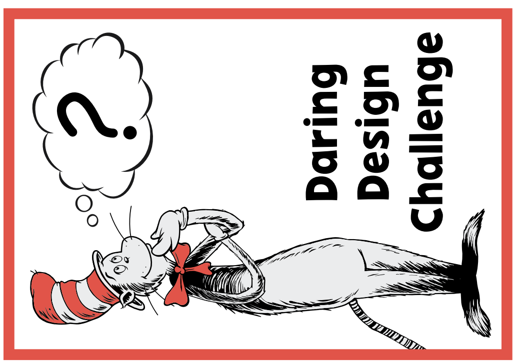


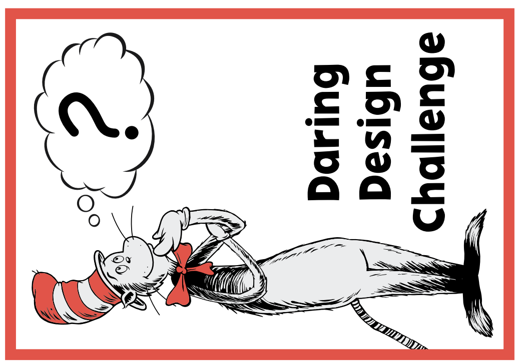












Create a musical instrument that can make three different notes or sounds.

Imagine & Plan!

Sounds can come from hitting or plucking an object, or blowing on or through objects of different lengths. What objects can help your design work?

Make a boat that floats.

Test!

Fill the bathtub or a sink with water to test your boat. Hope it floats!

Design something taller than you.

Create!

Have fun building together. Accidents happen, if you're lucky! Use a mistake to make it even better.

Build something that will balance a pencil.

Improve!

Can it balance two pencils? Three?

Design something that can launch a coin into the air.

Test!

How do you know if it is balanced?

Create something beautiful— just because.

Imagine & Plan!

Ask a friend how your creation makes him or her feel.

Make a ball that bounces.

Improve!

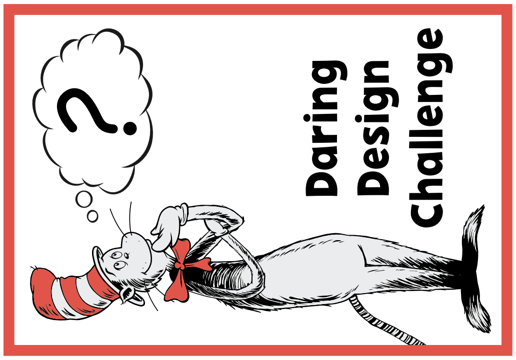
Can you make something that helps you move it faster?

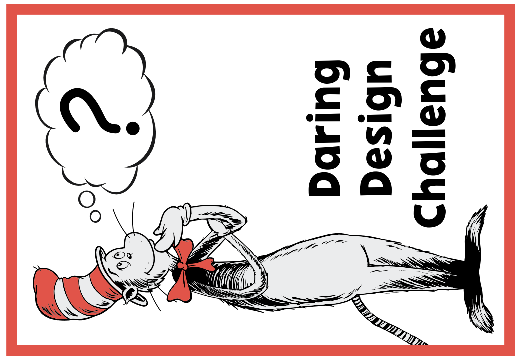
Build a tower as tall as your table using only paper.

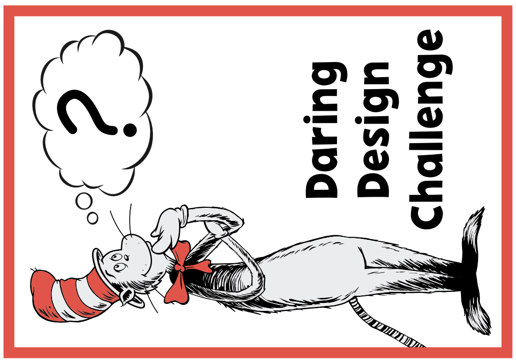
Imagine & Plan!

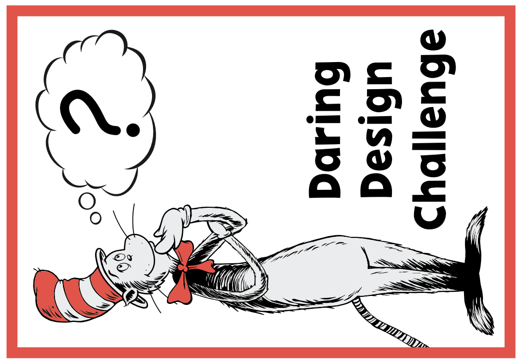
What can you do to make paper stronger and taller?

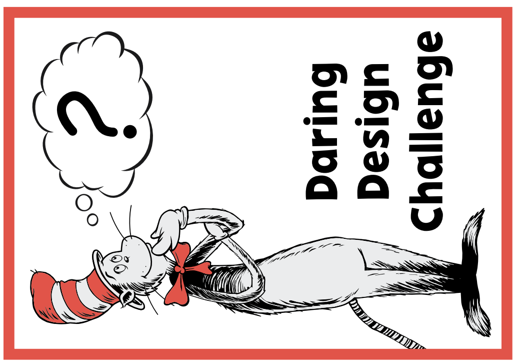
Print two-sided

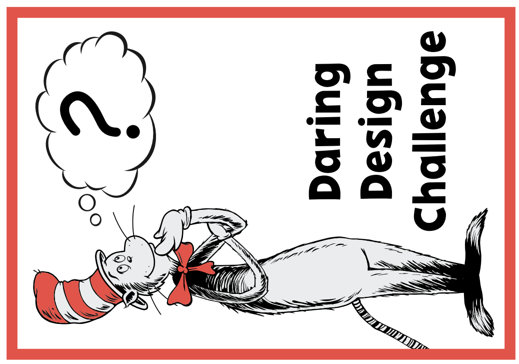


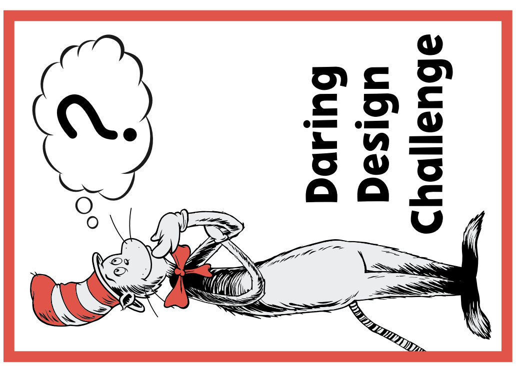


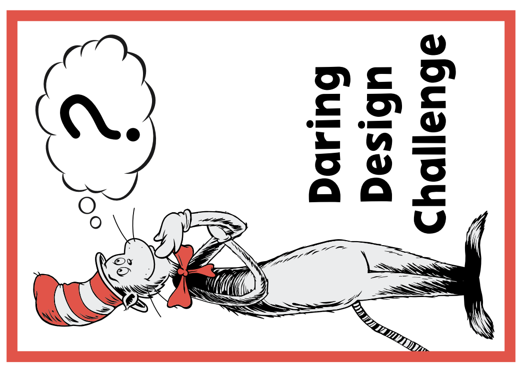












Build a shelter big and strong enough to protect a doll or stuffed animal.

Imagine & Plan!

What's the weather like where you live? Think of materials and objects that might keep away wind, cold, hot sun, rain, or snow!

Create a new sandwich.

Define the Problem!

Is there no jelly for your peanut butter? Or are you out of bread? Test your culinary creativity to find alternative solutions!

Create something with a handle to carry your design tools
(pencil, paper, tape, etc.)

Imagine & Plan!

Look at how many tools you have, and consider the size and how many compartments you might need.

Design a new game.

Define the Problem!

What if you and your friends need a new game to play outside or inside. Use a ball, dice, or whatever you imagine!

Measuring This and That

THE CAT IN THE HAT Knows a Lot About That!

Make and test The Cat in the Hat measuring tools.

**Which cereal box is taller? Are you taller than your friend?
Which tree trunk is wider? If you want to know for sure, you need to get evidence.
Make the Cat's hat and a Measuring Snail to measure and compare all sorts of things.**

Materials

- Cardstock or paper
- The Cat's Hat printable
- Measuring Snail printable

Create Your Measuring Tools

1. Using cardstock, print the Cat's hat and Measuring Snail printables.

   Tip No cardstock? No problem! Just glue the Cat's hat and the Measuring Snail to an old greeting card, flattened cereal box, or piece of cardboard for stability.

2. Cut out the Cat's hat and build the Measuring Snail as directed.
3. Use the stripes on the Cat's hat and the segments on your Measuring Snail's measuring line to measure the length of various household items.

Measure, Share, Compare!

1. Measure the length of the Cat's hat using the Measuring Snail. How long is it? Now measure the Measuring Snail with the Cat's hat. What do you notice?
2. Find 4 or 5 household items that are similar but different. (For example: different kinds of shoes, drinking glasses, toy cars, or chairs.) Predict which is the longest or shortest. Using the Cat's hat or the Measuring Snail, measure and compare the lengths. Did anything surprise you? Why?
3. What if you need to measure something round? What would be the best tool? Find a tree trunk, can of food, or something similar. Measure it using the Cat's hat and the Measuring Snail. What do you notice?
4. The Cat uses his hat to measure and improve the length of a stool leg when he visits Blueprintia. With the help of your measuring tools, design something that balances on three or four legs. Share your creation with a friend and use the Cat's hat or the Measuring Snail to show details about your design.

Find related games in the FREE *The Cat in the Hat Builds That* app. Download it now!

pbskids.org/catinthehat

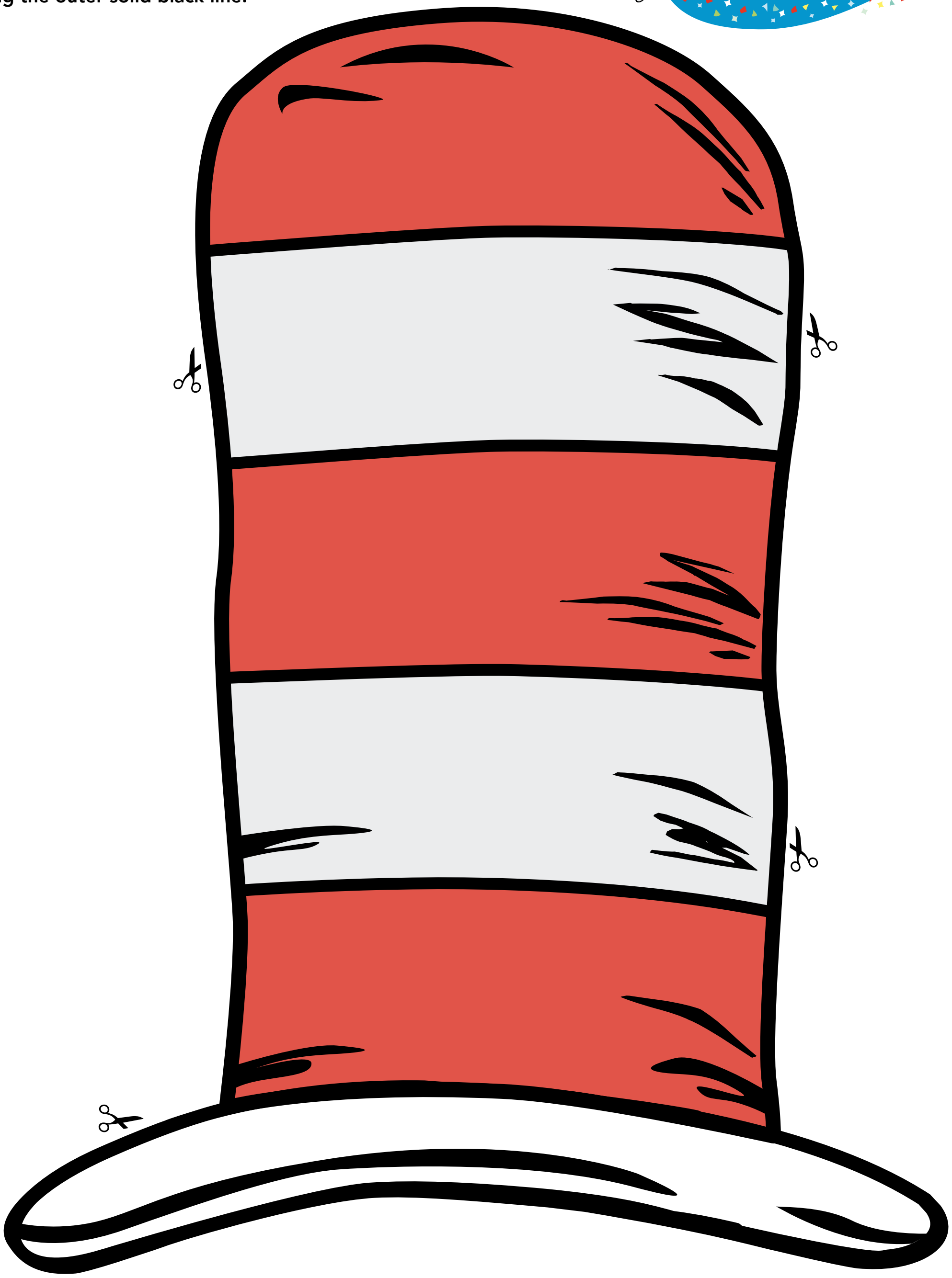
Powered by a Ready To Learn Grant

Corporation for Public Broadcasting

The contents of this document were developed under a cooperative agreement (PR/Award No. U295A150003, CFDA No. 84.295A) from the U.S. Department of Education. However, these contents do not necessarily represent the policy of the Department of Education, and you should not assume endorsement by the Federal Government.

The Cat's Hat

With an adult's help, cut out the Cat's hat along the outer solid black line.



The contents of this document were developed under a cooperative agreement (PR/Award No. U295A150003, CFDA No. 84.295A) from the U.S. Department of Education. However, these contents do not necessarily represent the policy of the Department of Education, and you should not assume endorsement by the Federal Government.

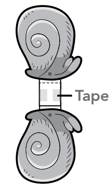
Measuring Snail

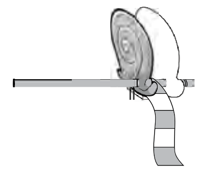
Materials

- Scissors
- Markers or crayons
- Clear tape
- Hole punch
- Unsharpened pencil

Instructions

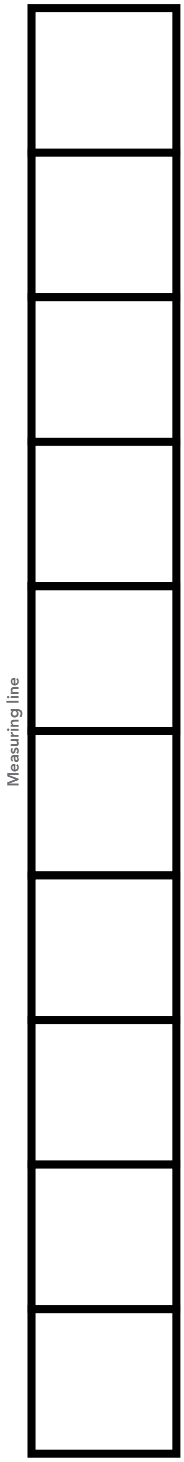
1. With an adult's help, cut out the two halves of the Measuring Snail's body and the measuring line.
2. Join the halves by overlapping the two tabs and taping them together.
3. Punch holes in each side of the Measuring Snail as marked.
4. Stick a pencil through the two holes.
5. Color the stripes on the measuring line. Use two colors and alternate them.
6. Tape one end of the measuring line—colors facing up—to the pencil, and roll up the measuring line by turning the pencil.
7. Pull the measuring line and put your Measuring Snail to work!





Test & Improve!

Explore the measuring ideas in the "Measuring This and That" activity. What could make your Measuring Snail better? Can you make it longer? What if you wanted it to pull easier and not bounce back? What other materials around your house could you use as a measuring line?



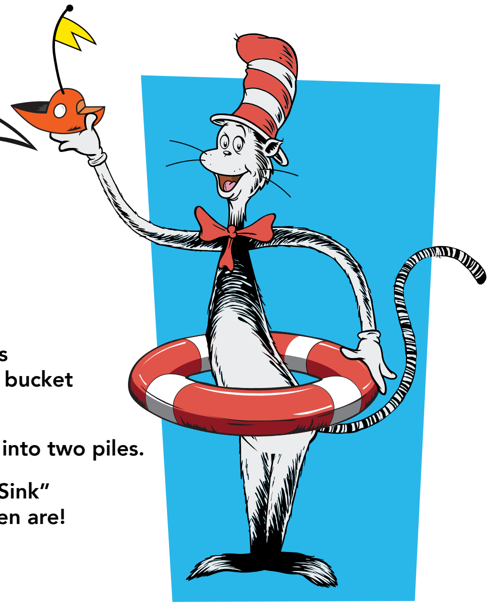
The contents of this document were developed under a cooperative agreement (PR/Award No. U295A150003, CFDA No. 84.295A) from the U.S. Department of Education. However, these contents do not necessarily represent the policy of the Department of Education, and you should not assume endorsement by the Federal Government.

What Floats Your Boat?

The CAT in the HAT Knows a Lot About That!

Predict and discover what sinks or floats!

During a trip to the Buoyant Sea, Sally and Nick test what sinks or floats. They learn that materials and shapes matter. What do you think will float? Test out your predictions using items from around the house and the printable Table.



Instructions

1. Print the Sink or Float Table (on the following page). If you are playing with a friend, print two.
2. With the help of an adult, gather ten small household items (rubber band, coin, pencil, sticker, etc.) and fill up a sink or bucket with water.
3. Predict whether the items will sink or float by sorting them into two piles.
4. Test your predictions in the water. Place each item under "Sink" or "Float" on the Table. Were you surprised? Scientists often are!

More Ways to Play

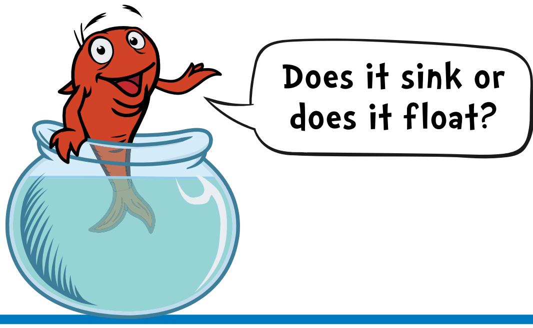
1. Can you change something that sinks into something that floats? Or something that floats into something that sinks? How did you do it?
2. At first, Sally and Nick think certain colors float. What do you think? Find ten, similar-colored items and predict whether they will sink or float. Now test and record your results with the Sink or Float Table. Does color matter? If not, what does?
3. After testing and placing your items on the Sink or Float Table, explain to a friend or family member why you think some items float and others sink.

pbskids.org/catinthehat

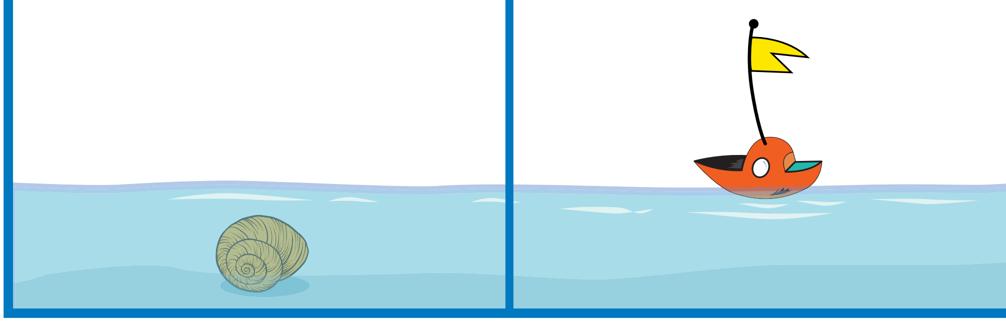
Powered by a Ready To Learn Grant

cpb Corporation for Public Broadcasting

The contents of this document were developed under a cooperative agreement (PR/Award No. U295A150003, CFDA No. 84.295A) from the U.S. Department of Education. However, these contents do not necessarily represent the policy of the Department of Education, and you should not assume endorsement by the Federal Government.



Sink	Float



Appendix D

Measure Development

All study instruments are available here: http://cct.edc.org/rtl/data-collection-tools.

Child Measures

LENS-Modified

The study team worked with the University of Miami team that developed the Lens on Science to create a modified version for use as a posttest assessment (hereafter referred to as Lens-Modified). We modified the full Lens assessment to align with the focal content domains in the *Cat in the Hat* resources. From the full pool of 498 Lens items, which include items designed to assess knowledge related to life science, earth and space science, physical science, and engineering and technology, the study team reviewed the 114 physical science items and the 43 engineering and technology items to determine which items were best aligned with the PBS KIDS Science Learning Framework[4] content domains covered in *The Cat in the Hat*. The study team then selected a subset of 60 items focused on physical science and engineering and technology. The study team selected items on the basis of the following criteria: (1) alignment with physical science and engineering core ideas targeted by the *Cat in the Hat* resources, (2) limiting the items' redundancy (for example, items that measured the same science core idea), and (3) the spread of items' difficulty levels to mirror the distribution of difficulty in the original assessment and to ensure an even distribution of items (i.e., not too easy or too difficult) to avoid possible floor and ceiling effects. The 60 selected items comprise 40 physical science items and 20 engineering and technology items. The resulting Lens-Modified assesses children's understanding of a broad set of physical science and engineering core ideas, inclusive of most of the physical science and engineering/technology core ideas covered in the *Cat in the Hat* resources. Each child was administered the same 60 items in random order using the "random mode" of the Lens software platform. See Table D1 and Figure D1 for a comparison of frequency and difficulty of the full set of Lens engineering/technology and physical science items and the Lens-Modified items.

4 As part of the development of science and engineering media under the Ready To Learn initiative, PBS developers created a learning framework, the PBS KIDS Science Learning Framework (Brenneman et al., 2018), with input from educational advisors to guide the development of media resources. The Framework is aligned with Head Start and Next Generation Science Standards (NGSS).

Table D1. **Distribution of the Difficulty of All Lens Physical Science & Engineering/ Technology Items Compared with Lens-Modified**

	All Items			Lens-Modified Items		
	n	Mean (*SD*)	Min, max	*n*	Mean (*SD*)	Min, max
Engineering/Technology	43	-0.50 (0.96)	-2.21, 1.61	20	0.43 (0.78)	-0.97, 1.61
Physical Science	114	0.25 (1.19)	-2.24, 3.28	40	0.91 (1.15)	-0.76, 3.28
Combined	157	-0.12 (1.15)	-2.48, 3.28,	60	0.62 (1.11)	-0.97, 3.28

Figure D1. **Distribution of the Difficulty of All Lens Physical Science & Engineering/ Technology Items Compared with Lens-Modified**

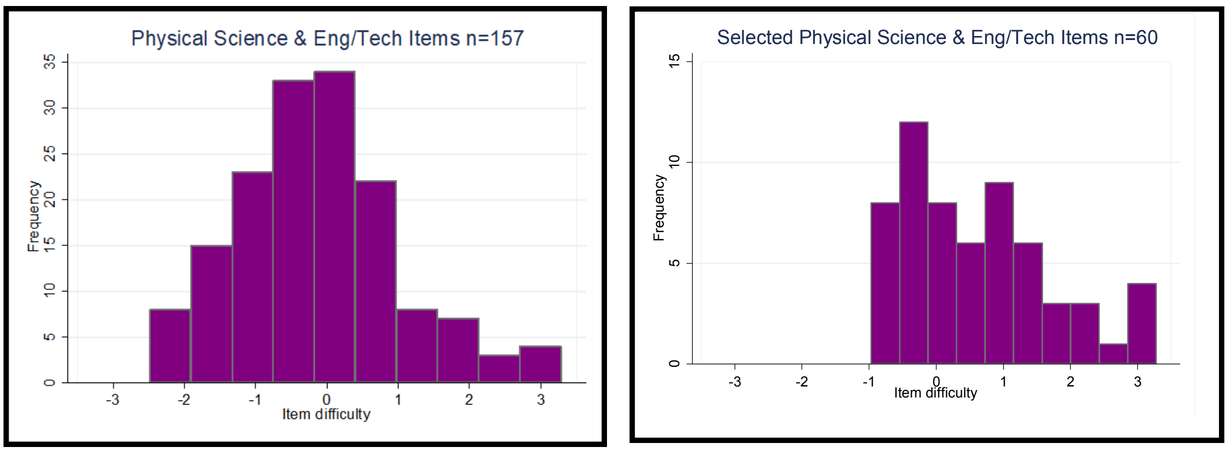


Figure D2. **Distribution of All Lens Physical Science Items Compared with Lens-Modified**

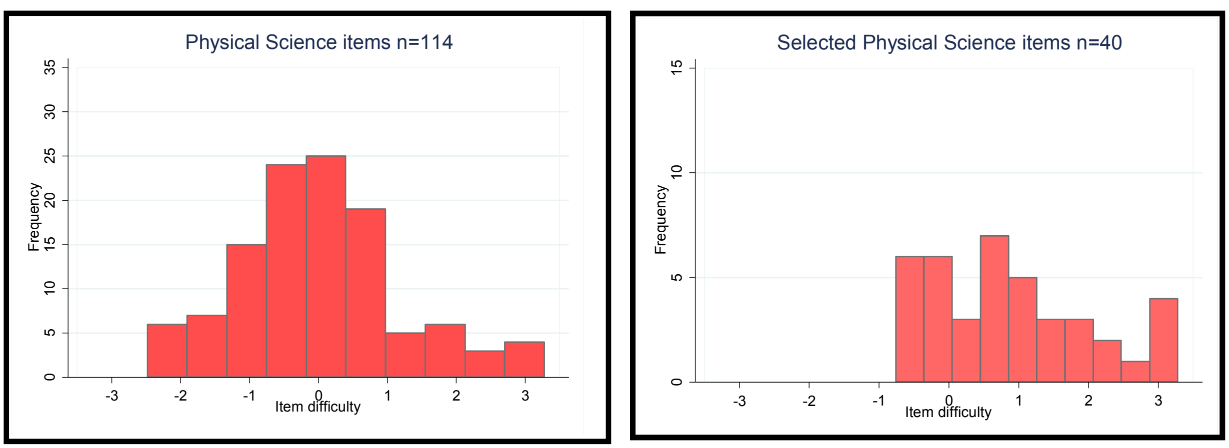
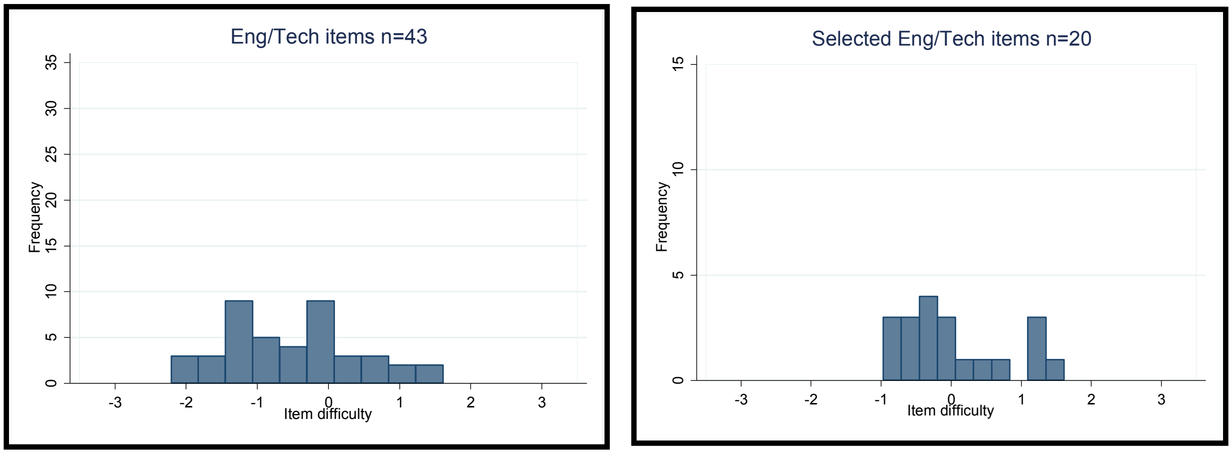


Figure D3. Distribution of All Lens Engineering/Technology Items Compared with Lens-Modified



Hands-On Preschool Assessments of Physical Science and Engineering

The study team developed three performance-based tasks to be administered at Meeting 2 to assess children's understanding of the role of properties (strength and length) and forces in structural stability, the role of properties (texture) and forces (friction) on movement down an incline, and how objects can be sorted on the basis of their material properties and uses. We developed this battery of tasks to complement the Lens-Modified, which covers a broad range of physical science and engineering/technology content, while the *Cat in the Hat* resources focus on a smaller subset of content. These tasks also tested children's ability to transfer learning from a digital environment to real-world and hands-on activities. Although *Cat in the Hat* games and videos address several physical science and engineering concepts and skills, the performance assessments focus on the three core concepts that received the most substantive focus in the games and videos and that were most closely aligned to the PBS KIDS Science and NGSS Frameworks.

Length, Strength, and Stability Task

The purpose of this task is to assess a child's understanding of how the properties of objects (such as size and shape) and materials (such as hardness and flexibility) make them suitable for different purposes and how different forces (pushes and pulls) can cause objects to move and influence the stability of a bridge. The assessor provides a child with objects of different "lengths" and "strengths" and asks the child to choose the most suitable object to use for the decking of a bridge that can support the weight of a toy character. Children have opportunities to use the Science and Engineering Practices as they plan and carry out investigations; construct explanations; engage in argument from evidence; and obtain, evaluate, and communicate information. This task aligns with the following PBS

KIDS Science Framework disciplinary core ideas and concepts within the Physical Science content domain:

PreK–K: Matter and Its Interactions

» Understand that different kinds of matter exist and that these can be described and classified by their observable properties.

» Understand that different properties of materials are suited for different purposes.

PreK–K Motion and Stability: Forces and Interactions

» Understand that pushes and pulls can cause objects to move.

» Understand that pushes and pulls can have different strengths and directions.

In addition, the task aligns with the following crosscutting concepts:

Crosscutting concepts

» **Stability and change.** By exploring and reflecting on various events and phenomena in the natural and human-designed world, children can observe that some things stay the same while other things change, and that things may change slowly or rapidly.

» **Cause and effect.** By identifying and observing events, relationships, and patterns in the natural and human-designed world, children can learn that events have causes that generate observable patterns, and they can design simple tests to gather evidence to support or refute their own ideas about causes.

» **Systems and system models.** By investigating accessible and visible systems in the natural and human-designed world, children can understand that objects and organisms can be described in terms of their parts. By describing things in terms of parts, roles of parts, and relationships among parts, children can understand that systems have parts that work together and that if a part of the system breaks, is removed, or is altered, the working of the system can change.

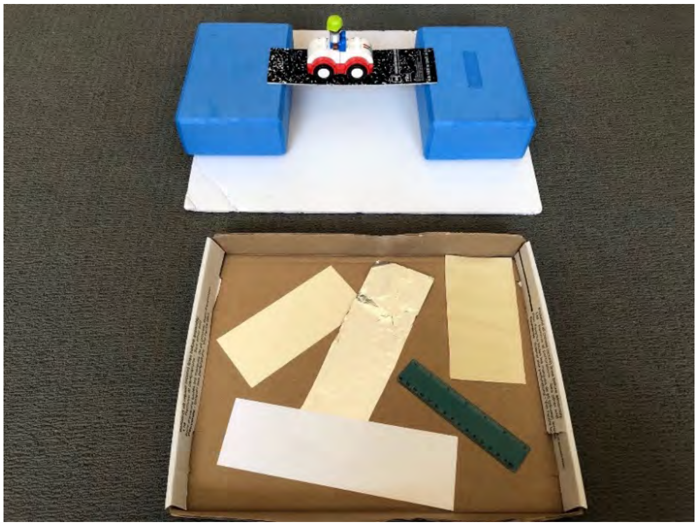
This assessment is designed to align with the learning goals of the *Cat in the Hat* theme Bridge-a-rama. Bridge-a-rama videos include episodes in which Nick, Sally, Cat, and other characters visit Spansylvania to explore what materials make a good bridge. The characters test objects made of different materials and different lengths (e.g., log, banana leaves, canoe) to try to cross a gap. They learn that an object must be both "long enough" and "strong enough" to make a successful bridge. In *The Cat in the Hat Builds That!* game Bridge-a-rama, the user recreates a similar scenario, choosing from a set of objects to test (e.g., log, stick, canoe) to find objects that are "long enough and strong enough." In the game, once users move the selected object(s) into the bridge position above the gap, they can select an icon to test the bridge. For some challenges, success is measured by whether the object is long enough to span the gap. For other challenges, success is measured by whether the object is strong enough, determined by how many birthday presents the dragon can successfully carry across the bridge. If the object is not long enough to span the gap, it falls down before a dragon tries to cross the bridge. If the object is not strong enough, the dragon drops some of his birthday presents on the way across.

Materials for the assessment include two yoga blocks, attached with Velcro to a foam poster board with an 8-inch gap between them; a Duplo character in a car; and the materials listed in Table D2 spread out in a large tray (Figure D4):

Table D2. Length, Strength, and Stability Materials

	Not strong enough	Strong enough
Not long enough	» Cardstock (6 inches)	» 6-inch ruler » Oven-ready lasagna noodles
Long enough	» Tin foil (9-inch by 2-inch strip) » Laminated piece of paper (9 inches by 2 inches)	» Composition notebook cover (cut in half lengthwise)

Figure D4. Length, Strength, and Stability Task Setup



The assessor begins by showing a child a set of objects spread out in a large tray that could be used to create a bridge so that a gender-neutral toy figure, "Sam," who is in a toy car on top of one block, can visit a friend on another block approximately 8 inches away. The assessor asks the child to select which object to use to help Sam get across to see her/his friend, and how s/he could try out that idea (i.e., test whether the object is long enough to span the gap between the blocks and strong enough to support Sam's weight). The assessor invites the child to test whether her/his choice worked. The child's score is based on whether s/he, within three attempts, selects the suitable object, tests each chosen object, and provides an explanation about why each tested object did or did not work, related to relevant properties and forces.

Scoring: Scores range from 0 to 8 points. The scoring rules for Length, Strength, and Stability are outlined in Table D3, and the levels of understanding demonstrated by the assessment are described in Table D4. See the rubric (Table D10) for details.

Table D3. Length, Strength, and Stability: How Learning was Measured and Scored in Relation to Physical Science Concepts and Science and Engineering Practices

Item #	Weight	PBS KIDS Science Framework Core Ideas, Concepts, Indicators, and Science & Engineering Practices		
		Matter and Its Interactions	**Motion and Stability: Forces and Interactions**	**Science and Engineering Practices**
A1-3	Up to 2 points (25%) 1 point for selecting the notebook piece, 0.5 points for selecting a material that is long enough.	Understand that different kinds of matter exist and can be described and classified by their observable properties. *Demonstrate an increased ability to observe, manipulate, describe, and ask questions about the characteristics and physical properties of familiar human-made and natural objects and liquids.* Understand that different properties of materials are suited for different purposes. *Distinguish between an object and the material from which it is made.* *Explore familiar objects to determine and describe how the materials of which they are made are related to the objects' properties.* *Identify the uses of various natural or human-made objects based on their properties.*	Understand that pushes and pulls can cause objects to move. *Explore and describe the effects of simple forces that push or pull in nature, such as wind and gravity.* Understand that pushes and pulls can have different strengths and directions. *Compare the effects of different strengths or different directions of pushes and pulls.*	Planning and Carrying Out Investigations *Make observations (firsthand or from media) to collect data that can be used to make comparisons and predictions and conclusions).*

Item #	Weight	PBS KIDS Science Framework Core Ideas, Concepts, Indicators, and Science & Engineering Practices		
		Matter and Its Interactions	**Motion and Stability: Forces and Interactions**	**Science and Engineering Practices**
B1-3	0 points*			Planning and Carrying Out Investigations *Make and test predictions based on background knowledge and prior experiences.*
C1-3	Up to 6 points (75%) 1-2 points for each time a correct explanation is given to "Why do you think [the object] did/ didn't work as a bridge?"	Understand that different properties of materials are suited for different purposes. *Distinguish between an object and the material from which it is made* *Identify the uses of various natural or human-made objects based on their properties.*	Understand that pushes and pulls can cause objects to move. *Observe and discuss ideas, based on evidence, about what makes something move, and how movements can be controlled and changed.*	Constructing Explanations *Use evidence to support a theory (science) or solution to a problem (engineering).* Engaging in Argument from Evidence *With support, compare results of an investigation with an original prediction and offer evidence as to why they do or do not match.* Obtaining, Evaluating, and Communicating Information *Share findings and explanations (correct or incorrect) with greater detail and through a variety of methods (e.g., telling adult or peer, writing/drawing in a journal).*

* Initially, the study team intended to award 1 point for each correct test of a bridge material (rolling the car across, placing the car in the center of the bridge, or holding an object up to the gap if the object was too short). Inspection of the data (conducted blind to condition), however, indicated that this effectively penalized children who selected longer objects, since it was easier to receive points for a short-object test. The study team elected not to award points from this item and more heavily weight the child's explanation of what happened (item C).

Table D4. **Levels of Understanding Demonstrated in the Length, Strength, and Stability Task**

Level of Understanding	Score Range
Strong understanding of how the properties of objects and materials as well as forces contribute to the stability of a bridge	6-8
Emerging understanding of how the properties of objects and materials as well as forces contribute to the stability of a bridge	3-5.5
No to weak understanding of the properties of objects and materials or forces and how they contribute to the stability of a bridge	0-2.5

Surfaces and Friction Task

The purpose of this task is to assess a child's understanding of how the properties of objects and materials and forces (including friction) influence the motion of objects. The assessor provides the child with three slides with differently textured (rough/smooth/sticky) surfaces and asks her/him to choose the slide that will enable a gender-neutral toy figure, "Sam," to slide down the fastest. Children have opportunities to use the Science and Engineering Practices as they plan and carry out investigations; analyze and interpret data; construct explanations; and obtain, evaluate, and communicate information. This task aligns with the following PBS KIDS Science Framework disciplinary core ideas and concepts within the Physical Science content domain:

PreK–K: Matter and Its Interactions

- Understand that different kinds of matter exist and that these can be described and classified by their observable properties.
- Understand that different properties of materials are suited to different purposes.

PreK–K Motion and Stability: Forces and Interactions

- Understand that pushes and pulls can cause objects to move.
- Understand that pushes and pulls can have different strengths and directions.

In addition, the task aligns with the following crosscutting concepts:

Crosscutting concepts

- **Structure and function.** By investigating how things work and reflecting on characteristic parts and what they do in both nature and the human-designed world, children can observe that the shape, material, and parts of an object or system are related to the function(s) of the object or system.
- **Cause and effect.** By identifying and observing events, relationships, and patterns in the natural and human-designed world, children can learn that events have causes that generate observable patterns, and they can design simple tests to gather evidence to support or refute their own ideas

about causes.

» **Systems and system models.** By investigating accessible and visible systems in the natural and human-designed world, children can understand that objects and organisms can be described in terms of their parts. By describing things in terms of parts, roles of parts, and relationships among parts, children can understand that systems have parts that work together and that if a part of the system breaks, is removed, or is altered, the working of the system can change.

This assessment is designed to align with the learning goals of the *Cat in the Hat* theme Slidea-ma-zoo. In this theme, children have the opportunity to watch and explore how the addition of materials (like melted butter, sand, honey, and ice) influence how fast the *Cat in the Hat* characters travel down slides. The addition of these materials changes the properties of the slides and influences the amount of friction generated, changing how fast the characters move down the slides.

Materials for the assessment include a plastic Duplo character and three wooden ramps (or slides). One ramp is covered with rubber, one with felt, and one with rough steel wool. The ramps consist of a 12-inch plank secured to a block. Each ramp is propped up at the same height to create ramps at equal inclines (Figure D3).

Figure D3. Surfaces and Friction Task Setup

During this assessment, the assessor presents the child with three slides that have the same incline but have different surface textures. Given that all variables other than texture are equal, Sam, the character in the task, will slide the fastest down the felt, more slowly down the rubber, and not at all down the steel wool. The assessor asks the child to look at and touch the slides and then to describe the texture of each. Subsequently, the assessor asks the child to describe how the slides are the same and different. With the slides set up side by side, the assessor then asks the child to (1) predict which slide Sam will slide down the fastest, (2) explain that prediction, and (3) state how the child might test on which slide Sam will slide the fastest. Finally, the assessor asks the child whether and how her/his prediction differs from what happened, and why s/he thinks so.

Scoring: Scores range from 0 to 14 points. The scoring rules for Surfaces and Friction are outlined in Table D5, and the levels of understanding demonstrated by the assessment are described in Table D6.

See the rubric (Table D11) for details.

Table D5. Surfaces and Friction: How Learning was Measured and Scored in Relation to Physical Science Concepts and Science and Engineering Practices

Item #	Weight	PBS KIDS Science Framework Core Ideas, Concepts, Indicators, and Science & Engineering Practices		
		Matter and Its Interactions	**Motion and Stability: Forces and Interactions**	**Science and Engineering Practices**
Part 1	Up to 6 points (43%) 1-2 points for accurate observation and description of texture for each of the three slides	Understand that different kinds of matter exist and can be described and classified by their observable properties. *Demonstrate an increased ability to observe, manipulate, describe and ask questions about the characteristics and physical properties of familiar human-made and natural objects and liquids.* Understand that different properties of materials are suited to different purposes. *Explore familiar objects to determine and describe how the materials of which they are made are related to the objects' properties.*		Planning and Carrying Out Investigations *Compare and contrast objects and events by describing similarities and differences in greater detail.* Obtaining, Evaluating, and Communicating Information *Describe observable phenomena using adjectives and labels.* *Use basic science and engineering content vocabulary when investigating and describing observable phenomena (e.g., mammal, life cycle, ecosystem, force).*
Part 2	Up to 4 points (29%) 1-2 points for an accurate observation about character-istic/s that are the same about the three slides and 1-2 points for an accurate observation about character-istics that are different	Understand that different kinds of matter exist and can be described and classified by their observable properties. *Demonstrate an increased ability to observe, manipulate, describe, and ask questions about the characteristics and physical properties of familiar human-made and natural objects and liquids.*		Planning and Carrying Out Investigations *Compare and contrast objects and events by describing similarities and differences in greater detail.* Obtaining, Evaluating, and Communicating information *Describe observable phenomena using adjectives and labels.*

Item #	Weight	PBS KIDS Science Framework Core Ideas, Concepts, Indicators, and Science & Engineering Practices		
		Matter and Its Interactions	**Motion and Stability: Forces and Interactions**	**Science and Engineering Practices**
Part 3	Up to 2 points (14%) 1-2 points for justifying their prediction of which slide Sam will slide fastest on	Understand that different properties of materials are suited for different purposes. *Distinguish between an object and the material from which it is made.* *Identify the uses of various natural or human-made objects based on their properties.*	Understand that pushes and pulls can cause objects to move. *Observe and discuss ideas, based on evidence, about what makes something move, and how movements can be controlled and changed.* *Plan and conduct an investigation to compare different types of pushes and pulls.* Understand that pushes and pulls can have different strengths and directions. *Compare the effects of different strengths or different directions of pushes and pulls.*	Planning and Carrying Out Investigations *Demonstrate a greater ability to make (and test) predictions based on background knowledge and experiences.*
Part 4	0 points*	Understand that different properties of materials are suited for different purposes. *Distinguish between an object and the material from which it is made.* *Identify the uses of various natural or human-made objects based on their properties.*	Understand that pushes and pulls can cause objects to move. *Explore and describe the effects of simple forces that push and pull in nature such as wind, gravity.* *Observe and discuss ideas, based on evidence, about what makes something move, and how movements can be controlled and changed.* Understand that pushes and pulls can have different strengths and directions. *Compare the effects of different strengths or different directions of pushes and pulls.*	Planning and Carrying Out Investigations *Demonstrate a greater ability to (make and) test predictions based on background knowledge and experiences.* *With guidance, plan and conduct simple investigations using simple tools (i.e., articulate steps to be taken and materials to use for exploring testable questions).*

Item #	Weight	PBS KIDS Science Framework Core Ideas, Concepts, Indicators, and Science & Engineering Practices		
		Matter and Its Interactions	**Motion and Stability: Forces and Interactions**	**Science and Engineering Practices**
Part 5	Not scored			Analyzing and Interpreting Data *Talk about and reflect on what happened during an investigation (or when solving a problem or why it solved a problem).*
Part 6	Up to 2 points (14%) 1-2 points for accurate explanation of why Sam slid the fastest on that particular slide	Understand that different properties of materials are suited for different purposes. *Distinguish between an object and the material from which it is made.* *Identify the uses of various natural or human-made objects based on their properties.*	Understand that pushes and pulls can cause objects to move *Observe and discuss ideas, based on evidence, about what makes something move, and how movements can be controlled and changed.* Understand that pushes and pulls can have different strengths and directions. *Compare the effects of different strengths or different directions of pushes and pulls.*	Constructing Explanations *Use evidence to support a theory (science) or solution to a problem (engineering).* Obtaining, Evaluating, and Communicating Information *Share findings and explanations (correct or incorrect) with greater detail and through a variety of methods (e.g., telling adult or peer, writing/drawing in a journal).*

* Initially, the study team intended to award up to 2 points for describing how to test which slide was the fastest (sliding an object down 1-2 slides for 1 point and down all 3 slides for 2 points) and also 1 point for correctly identifying whether or not their prediction matched what happened. Analysis of item characteristics (conducted blind to condition), however, showed that these two items did not correlate with the other items in the task, indicating that they may not be measuring the same intended construct. As a result, these items were not included in the final calculation of scores.

Table D6. **Levels of Understanding Demonstrated in the Surfaces and Friction Task**

Level of Understanding	Score Range
Strong understanding of the properties of objects and materials and the forces (pushes and pulls, including friction) that influence how objects move	8-14
Emerging understanding of the properties of objects and materials and the forces (pushes and pulls, including friction) that influence how objects move	4-7
No to weak understanding of the properties of objects and materials and the forces (pushes and pulls, including friction) that influence how objects move	0-3

Colors, Shapes, and Uses Task

The goal of this task is to assess a child's understanding of sorting, and that different objects can be described and categorized on the basis of their observable properties or functions. Children identify similarities and differences among colors of objects, sort objects on the basis of shape (with picture cues), complete a sort based on use, and fix a sort based on color. Children have the opportunity to use the Science and Engineering Practice: Mathematics and Computational Thinking. This task aligns with the following PBS KIDS Science Framework disciplinary core idea and concepts within the Physical Science content domain:

PreK–K: Matter and Its Interactions

» Understand that different kinds of matter exist and that these can be described and classified by their observable properties.

» Understand that different properties of materials are suited to different purposes.

In addition, the task aligns with the following crosscutting concepts:

Crosscutting concepts

» **Structure and function.** Solid, nonporous materials like metal or plastic are good materials for use as plates and bowls; round objects like balls move by rolling.

This assessment is designed to align with the learning goals of the *Cat in the Hat* theme Sorta-ma-gogo. In this theme, Nick and Sally go to "Tomorrow Land" with the *Cat in the Hat* and learn about how objects can be sorted in different ways (by color, shape, texture, and use). In the related Sorta-ma-gogo game, the player has opportunities to collect objects that are similar in one property or use (same color, same shape, same use) and then sort objects on the basis of similarities and differences in color, shape, texture, and use. Subsequent levels increase in difficulty as distractor items are added to the set and less salient characteristics are used as the basis for sorting (for example, texture).

Materials for the assessment include 21 objects incorporating a variety of colors, shapes, and uses (Table D7 and Figure D4); 3 plastic trays; 3 shape cards (circle, square, triangle); and 3 picture cards with images of trays containing the following objects: *Tray 1:* red cup, red marker, red block, blue Cookie Monster; *Tray 2:* orange felt, orange spoon, orange ball; *Tray 3:* blue fork, blue ball, blue streamer roll.

Table D7. Colors, Shapes, and Uses Task Materials

Play	Art	Eat
Blue cube block	Yellow felt triangle	Round orange (artificial fruit)
Red triangle block	Red felt	Square blue napkin
Blue round ball	Red triangle crayon	Round red plate
Orange ping pong ball	Blue colored pencil	Blue fork
Blue Cookie Monster character	Orange marker	Orange spoon
Red round ball	Orange square notepad	Triangular pizza slice (artificial)
Orange Ernie character	White felt triangle	
Natural wood triangle block		

Figure D4. Colors, Shapes, and Uses Task Setup

Scoring: Scores range from 0 to 8 points. The scoring rules for Colors, Shapes, and Uses are outlined in Table D8, and the levels of understanding demonstrated by the assessment are described in Table D9. See the rubric (Table D12) for details.

Table D8. **Colors, Shapes, and Uses: How Learning was Measured and Scored in Relation to Physical Science Concepts and Science and Engineering Practices**

Item #	Weight	PBS KIDS Science Framework Core Ideas, Concepts, Indicators, and Science & Engineering Practices	
		Matter and Its Interactions	**Science and Engineering Practices**
1A-1D	Up to 4 points (50%)	Understand that different kinds of matter exist and can be described and classified by their observable properties. *Demonstrate an increased ability to observe, manipulate, describe, and ask questions about the characteristics and physical properties of familiar human-made and natural objects and liquids.*	Using Mathematics and Computational Thinking *Sort and categorize observable phenomena on the basis of attributes such as appearance, weight, function, ability, texture, odor, and sound.*
2A	Up to 1 point (12.5%)	Understand that different kinds of matter exist and can be described and classified by their observable properties. *Demonstrate an increased ability to observe, manipulate, describe, and ask questions about the characteristics and physical properties of familiar human-made and natural objects and liquids.*	Using Mathematics and Computational Thinking *Sort and categorize observable phenomena on the basis of attributes such as appearance, weight, function, ability, texture, odor, and sound.*
2B	Up to 1 point 12.5%) (If a child requires additional prompting, maximum points awarded are 0.5).	Understand that different kinds of matter exist and can be described and classified by their observable properties. *Demonstrate an increased ability to observe, manipulate, describe, and ask questions about the characteristics and physical properties of familiar human-made and natural objects and liquids.*	Using Mathematics and Computational Thinking *Sort and categorize observable phenomena on the basis of attributes such as appearance, weight, function, ability, texture, odor, and sound.*
3A-3C	Up to 1 point (12.5%)	Understand that different properties of materials are suited to different purposes. *Identify the uses of various natural or human-made objects on the basis of their properties.*	Using Mathematics and Computational Thinking *Sort and categorize observable phenomena on the basis of attributes such as appearance, weight, function, ability, texture, odor, and sound.* Obtaining, Evaluating, and Communicating Information *Share findings and explanations (correct or incorrect) with greater detail and through a variety of methods.*

Item #	Weight	PBS KIDS Science Framework Core Ideas, Concepts, Indicators, and Science & Engineering Practices	
		Matter and Its Interactions	**Science and Engineering Practices**
4A-4C	Up to 1 point (12.5%)	Understand that different kinds of matter exist and can be described and classified by their observable properties. *Demonstrate an increased ability to observe, manipulate, describe, and ask questions about the characteristics and physical properties of familiar human-made and natural objects and liquids.*	Planning and Carrying Out Investigations *Compare and contrast objects and events by describing similarities and differences in greater detail.* Using Mathematics and Computational Thinking *Sort and categorize observable phenomena on the basis of attributes such as appearance, weight, function, ability, texture, odor, and sound.* Obtaining, Evaluating, and Communicating Information *Share findings and explanations (correct or incorrect) with greater detail and through a variety of methods.*

Table D9. Levels of Understanding Demonstrated in the Colors, Shapes, and Uses Task

Level of Understanding	Score Range
Strong understanding that objects can be sorted on the basis of their properties and/or functions	5.6-8
Emerging understanding that objects can be sorted on the basis of their properties and/or functions	3.1-5.5
No to weak understanding that objects can be sorted on the basis of their properties and/or functions	0-3

Table D10. **Length, Strength, and Stability Rubric**

Prompt	General Description of Correct Response	Examples of High Quality Responses (1 point)	Examples of Mid Quality Responses (0.5 point)	Examples of Incorrect Responses (0 points)	Total points possible
A. Object: Which object could you use to help Sam get from this block? (administered 3 times)	Choice of the material that was both long enough and strong enough	chose notebook piece one or more times	Choice of other material that was long enough (foil or laminated paper). 0.5 points for each, up to 1 point	never chose notebook piece	2
		Examples of Correct Responses		**Incorrect Responses**	
B. Test: How can you try out your idea? (administered 3 times)	Accurate test of object as a bridge	Foil, laminated paper: » "Rolled car across" » "Placed car in center of bridge" » "Other" acceptable response for foil: child did not test using the car because it was clear it was not going to work, e.g. if child placed the foil appropriately and it did not stay up Lasagna, cardstock, or ruler: » "Held object up, object was too short" » "Other" acceptable response: place on one block, showing that it doesn't reach; held object up and also attempted to roll across and/or support object with hand Notebook piece: » "Rolled car across" » "Placed car in center of bridge" (or somewhere on bridge) » "Supported object or car with hand" (decided based on inconsistency in assessor scoring)"		Lasagna, cardstock, ruler, notebook piece: » "Did not test" Foil, laminated paper: » "Did not test" » "Supported object or car with hands" (selected alone or in combination with another option)"	0. No points included in final score based on final analyses.

Prompt	General Description of Correct Response	Examples of High Quality Responses (2 points): precise and accurate explanation related to a property of the object, not what the object is doing	Examples of Mid Quality Responses (1 point): less precise but accurate explanation related to a property of the object (small vs. short), or an accurate explanation related to the configuration of the object in relation to the blocks (e.g. stable)	Examples of Incorrect Responses (0 points): inaccurate explanation related to a property of the object (e.g. ruler was too weak), reiteration that the bridge worked (e.g. because it didn't break), explanation related to a property of the car (e.g. the car was too heavy)	Total points possible
C. Explanation: Why do you think it did/didn't work as a bridge? (administered 3 times)	Accurate explanations related to a property of the object (not what the object is doing)	Lasagna/Ruler/Cardstock: » too short » not long enough Laminated paper/Foil: » not strong enough » too bendy/it can bend/it is bending » too weak » not sturdy » not hard enough » too soft » wobbly Notebook piece: » strong » hard » long » not bendy/it doesn't bend as much » sturdy	Lasagna/Ruler/Cardstock: » too small » too little » too tiny » not tall enough » it doesn't reach/it has to be this long/not far enough Laminated paper/Foil: » not thick » thin » not stable Foil: » it fell/it is falling Notebook piece: » big enough/tall enough » not too small » thick enough » not too thin » stable/stabilized	All materials: » because it worked » because Sam fell down » because it's made of (material) » because the car/Sam was too heavy Lasagna/Ruler/Cardstock: » too weak Laminated paper/Foil: » too short » not heavy enough Notebook: » because it won't break » because it won't fall	6
				Score Range:	**0-8**

Table D11. Surfaces and Friction Rubric

Prompt	General Description of Correct Response	Examples of High Quality Responses (2 points): accurate adjective or descriptive word, non-real words are ok	Examples of Mid Quality Responses (1 point): "like a" accurate description, accurate but negative textural description, accurate description of feeling but not about texture	Examples of Incorrect Responses (0 points): inaccurate description, "like a" inaccurate description	Total Points Possible
1a. Rubber—How does it feel?	accurate observation about a textural characteristic of the rubber slide*	smooth, sticky, slippery	cold, flat, hard, not rough, not fluffy	fast, slow, soft, slidey	2
1b. Felt—How does it feel?	accurate observation about a textural characteristic of the felt slide*	soft, fluffy, fuzzy, furry, smooth	hard, comfy	fast, slow, different, rough, good, bumpy, scratchy, light, grassy, like a sponge	2
1c. Steel Wool—How does it feel?	accurate observation about a textural characteristic of the steel wool slide*	rough, scratchy, bumpy, spikey, pokey	hard, grassy, like grass, itchy, tickly, not smooth, fuzzy	fast, slow, different, crumply, clunky	2
	**the inclusion of a correct observation gets a point even if it's coupled with an incorrect observation as long as they're not clearly contradictory; not real words are ok, non-adjective descriptions are ok*				

Prompt	General Description of Correct Response	Examples of High Quality Responses (2 points): accurate identification of characteristic that is the same AND related to how an object would slide down	Examples of Mid Quality Responses (1 point): accurate identification of characteristic that is the same BUT NOT related to how an object would slide down	Examples of Incorrect Responses (0 points): identification of a characteristic that is not actually the same, or an answer that is too vague, or identification of something that is the same about only 2 slides	Total Points Possible
2a. What is the same about the slides?	accurate observation about characteristic/s that are the same	Incline/Steepness, Size (Length/Width/Height), shape » all have stuff on them	all have small, plain strip at the bottom of each? Ie. they're all "tan at the bottom?" » all have staples » all have wooden base/have block on back » all are on the table	desription of differences between slides, they look the same, beautiful, they slide, these two are the same color	2

Prompt	General Description of Correct Response	Examples of High Quality Responses (2 points): accurate identification of characteristic that is different AND related to how an object would slide down	Examples of Mid Quality Responses (1 point): accurate identification of characteristic that is different BUT NOT related to how an object would slide down OR a description of all three that is less than perfectly accurate	Examples of Incorrect Responses (0 points): identification of a characteristic that is not actually different, or an answer that is too vague	Total Points Possible
2b. What is different about the slides?	accurate observation about characteristic/s that are different; if they describe individual slides, they must describe all three	» they feel different » this one is smooth, this one is soft, and this one is scratchy » there are different things on them	Color (name different colors or say "color") » this one is hard, this one is hard, and this one is soft	Incline/Steepness Size (Length/Width/Height), description of a characteristic of just one slide, slow, fast » does not describe all three	2

Prompt	Description of Correct Response	Examples of High Quality Responses (2 points): identifies a physical characteristic related to how things might move on the slide, and it's accurate	Examples of Mid Quality Responses (1 point): not completely accurate physical characteristic	Examples of Incorrect Responses (0 points)	Total Points Possible
3a. Do you think Sam should use this slide, this slide, or this slide?	n/a; child predicts either rubber, felt, or steel wool slide.	n/a	n/a	n/a	Not scored—team did not plan to score this.
3b. Why will Sam go fastest on this slide?	justification of prediction; identifies a physical characteristic related to why an object might move faster, or a characteristic of the two other slides that attempt to explain why they were slower	» because it's softer, smooth, flat » because these ones are rougher (referencing slide that was not chosen)	rough, stable, solid, strong, hard	because it's faster/feels fast, because he likes it more, color (because it's black), because it looks like a road	2
Prompt	**General Description of Correct Response**	**Examples of High Quality Responses**	**Examples of Mid Quality Responses**	**Examples of Incorrect Responses**	**Total Points Possible**
4. How could we test?	description of a valid test to determine speed; can include verbal suggestions or gestural indications that a valid test would involve sliding Sam down all three slides	» Three slides (whether re-prompting was done or not » Child slides finger down each slide.	» One slide » Two slides » Slide down this one.	No response, or not related to sliding	0. No points included in final score based on final analyses.

Prompt	General Description of Correct Response	Examples of High Quality Responses	Examples of Mid Quality Responses	Examples of Incorrect Responses	Total Points Possible
5. On which slide did Sam go the fastest?	n/a; child identifies the slide on which Sam went fastest.	n/a	n/a	n/a	Not scored—team did not plan to score this.
6a. Did what happened to Sam match your prediction?	accurate identification of whether or not their prediction matched the result of the test	n/a	» Yes they did match » "it did match and it did not match" (if two were same speed)"	"This one was fastest" (doesn't respond to the question about whether the prediction was right or wrong)	0. No points included in final score based on final analyses.
Prompt	**General Description of Correct Response**	**Examples of High Quality Responses (2 points):** identifies a physical characteristic related to how things might move on the slide, and it's accurate	**Examples of Mid Quality Responses (1 point):** not completely accurate physical characteristic, a negative textural description	**Examples of Incorrect Responses (0 points):** related to something constant across the slides, or a characteristic not related to sliding	**Total Points Possible**
6b. Why did Sam go fastest on this slide?	accurate explanation of why Sam slid down the fastest on that particular slide; must accurately reference textural characteristics of the felt slide or accurately reference textural characteristics of the other slides	soft/softer, slippery/more slippery, smooth/smoother, fuzzier (in reference to felt) OR the other slides were too rough/bumpy/sticky (in reference to the rubber and steel wool slides)	harder, not bumpy	Incline/steepness, size, color, it's cold"	2
				Score Range:	**0-14**

Table D12. **Colors, Shapes, and Uses Rubric**

Prompt	General Description of Correct Response	Examples of Correct Responses	Examples of Incorrect Responses	Total Points Possible
1a. Can you find 3 things that are different colors?	identifies at least 3 objects that are different colors, and no repeat colors	1 each of red, blue, orange, yellow, or white object; ok to choose 4 or 5 as long as they are unique In general consider pizza to be orange, cookie to be blue, and Ernie to be orange, unless a child ID it as another color correctly If a child calls something a debatable color (e.g. ping pong ball as yellow), then we consider based on what the child says.	only 2 objects of different colors 1 multi-colored object (e.g. pizza) including multiple objects of the same color (e.g. 2 red things, 2 blue, 2 orange)	1
1b. Can you find 3 things that are the same color?	identifies at least 3 objects of the same color, and no objects of a different color	3, 4, or 5 red, blue, or orange objects (pizza counts as orange unless child states otherwise)	only 2 objects of the same color 3 objects of the same color, and 1 of a different color	1
1c. Can you find 3 things we could use for eating?	identifies at least 3 eating objects	3 of the following: napkin, plate, fork, spoon, pizza, orange Other objects can be acceptable if a child explains what they are using it as (e.g. yellow felt as cheese or napkin, white felt as napkin)	only 2 objects for eating	1
1d. Can you find 3 things we could use for art?	identifies at least 3 art objects	3 of the following: pencil, crayon, marker, paper, 1 of the felt pieces Other objects can be acceptable if a child explains what they are using it as (e.g. napkin to draw on)	only 2 objects for art	1
2a. What do you think the word "sort" means?				1

Prompt	General Description of Correct Response	Examples of Correct Responses	Examples of Incorrect Responses	Total Points Possible
	something about same/similar things together AND different things in other groups	e.g. put the same things together in one group and different things in a different group		1 point
	something about same/similar things together OR different things in other groups	e.g. put the same things together		.8 points
	gives a specific example of a way to sort	e.g. get all the blue things, which colors go together, matching colors		.6 points
	says something about putting things in piles/groups, without mentioning a characteristic	e.g. put things in groups, separating things		.4 points
	says something about potential sorting categories without saying that you put them in groups	e.g. clean and dirty clothes, sorting red and blue, big and small, bottles and cans		.4 points
	says something about same/ different but doesn't say anything about putting in groups or characteristics	e.g. some same some different, matching		.4 points
	uses "sort/sorting" in the definition and references some characteristic that is used for sorting	e.g. sorting colors/garbage/toys		.2 points
	says something about cleaning, organizing	e.g. put things in other things, clean up your toys, put things where they go, organizing		.2 points

Prompt	General Description of Correct Response	Examples of Correct Responses	Examples of Incorrect Responses	Total Points Possible
2b. …sort as many of these things as you can by shape	1/14 point for each correctly sorted item, up to 14; 1/14 point deducted for each incorrectly sorted item with the following exceptions: circle tray: marker, pencil, cookie monster, spoon (no points awarded or deducted) triangle tray: crayon (no points awarded or deducted) If there are more incorrect items than correct items sorted, points will not be negative; the lowest possible score is zero	circle tray: plastic orange, red plate, blue ball, red ball, ping pong ball square tray: blue block, napkin, red felt, orange paper triangle tray: white felt triangle, yellow felt triangle, pizza slice, plain triangle block, red triangle block"	circle tray: pizza square tray: red ball triangle tray: napkin	1
2b. FOLLOW UP…sort as many of these things as you can by shape (scored only if child does not respond to first prompt and assessor needs to model sorting of three objects)	1/28 points for each correctly sorted item, up to 11 objects; 1/28 points deducted for each incorrectly sorted item with the following exceptions: » circle tray: marker, pencil, cookie monster, spoon (no points awarded or deducted) » triangle tray: crayon (no points awarded or deducted) If there are more incorrect items than correct items sorted, points will not be negative; the lowest possible score is zero	circle tray: red plate, blue ball, red ball, ping pong ball square tray: napkin, red felt, orange paper triangle tray: yellow felt triangle, pizza slice, plain triangle block, red triangle block	circle tray: pizza square tray: red ball triangle tray: napkin"	0.5

Prompt	General Description of Correct Response	Examples of Correct Responses	Examples of Incorrect Responses	Total Points Possible
3c. Why did you put the spoon in that tray?	if child correctly places all three objects in trays, provides accurate explanation for all 3 (for spoon, says something related to eating, food; for marker, says something related to art, making stuff, writing; for block, says something related to toys, playing)	for spoon: food things, kitchen things	Incorrectly placed 1-3 objects; Correctly explained 1 or 2 of the 3 objects; Incorrect explanations: for spoon - because it goes with the fork, because it belongs/goes there, because it's orange/round/etc. for marker - because it goes with the crayon, because you can color on the paper, because it belongs/goes there for block - it's for building, because there's another block in there, because it belongs/ goes there	1
4a. How do you think my friend was trying to sort these things?	1/3 point for saying something related to color	by color: blue, orange, red	they did it wrong, He should have put Cookie Monster with the blue things	1
4b. Which object do you think is in the wrong tray?	1/3 point for identifying cookie monster	points to or says "cookie monster"		
4c. Which tray would be a better place for Cookie Monster?	1/3 point for identifying the blue tray	points to or says "blue/blue tray"		
			Score Range:	**0-8**

Hands-on Task Psychometric Analyses

The scale developers created draft rubrics for all items in each scale. On the basis of inspection of an initial subset of data, we refined these rubrics to include more detailed examples of correct and incorrect responses. For all three scales, two coders individually coded half of all open-ended child responses and double-coded an additional 20% of the completed assessments (58 assessments for each scale). For the sake of consistency across tasks, one common coder coded all three tasks. We evaluated interrater reliability by using weighted kappas. After each coder had completed her scoring, the two coders reviewed discrepant codes and reached consensus on a final code.

To assess item characteristics and internal consistency, we inspected means and distributions of all items, Cohen's alpha, corrected item-total correlations, and confirmatory factor analysis (CFA) results when applicable. On the basis of these results, we did not include some items in final total scores. We summarize these results for each scale below.

Length, Strength, and Stability

For Length, Strength, and Stability (see Table D10), coders scored children's explanations of why a bridge did or did not work (Item C) with possible scores of 0, 1, or 2. Interrater reliability for these items was high, with an average weighted kappa of .91 across the three iterations of Item C. Items that involved more concrete responses (Item A: choosing a material; Item B: how the child tested the material) were scored by one coder according to scoring rules; any questionable responses that did not align clearly with scoring rules were scored by the two coders by consensus.

Internal consistency was first evaluated for the original scale, including Item A, three iterations of Item B, and three iterations of Item C. Cronbach's alpha was low at .486, although this was not entirely surprising considering the small number of items. Corrected item-total correlations for the three iterations of Item B were very low, ranging from .095 to .135. CFA also indicated poor model fit, $X^2(14, N = 287) = 69.4$, $p < .001$, CFI = .62, RMSEA = .12, and relatively low factor loadings of the three iterations of Item B (less than .4). We further inspected Item B scores by comparing mean scores when children chose shorter materials (lasagna noodle, cardstock, or ruler) compared with longer, bendy materials (foil, laminated paper). Children were significantly more likely to get this item correct when they chose the shorter materials; in other words, when children chose a shorter material, they received a substantially easier item. There were also some discrepancies reported by assessors in the way they recorded child responses to Item B when children chose the notebook piece; it was difficult for assessors to distinguish if children were fairly testing this material by rolling the car across the bridge or if they were slightly supporting the car across the bridge with their hand as they moved it. For these reasons, we decided not to include the scores for Item B in the final score for this scale. After removing these items, Cronbach's alpha did not increase substantially (.489); but considering that it was based on only four items, the fact that it did not decrease was encouraging.

Children were also more likely to score higher on Item C if they chose a shorter material; it was easier for children to explain that a bridge didn't work because it was "too short" rather than because it was "too bendy" or "not hard enough." This difference was less dramatic, however, and we decided to offset this difference by awarding a half point for choosing a longer, bendy material up to two times. The final resulting scale, using a sum score, had a normal distribution (range from 0 to 8, mean of 4.27, standard deviation of 1.96) with no indication of skewness or kurtosis (0.11 and -1.05, respectively).

Surfaces and Friction

For Surfaces and Friction (see Table D11), coders scored the following open-ended items with possible scores of 0, 1, or 2: children's descriptions of the three materials (Items 1a, 1b, 1c), observations of similarities and differences between the slides (Items 2a, 2b), justification of their prediction of the fastest slide (Item 3b), and explanation of which slide was the fastest (Item 6b). Interrater reliability for these items was very good, with an average weighted kappa of .86 across the seven items. Items that involved more concrete responses (Item 4: child's indication how to test which slide was fastest; Item 6a: child's indication of whether or not the result of the test matched their prediction) were scored by one coder according to scoring rules; any questionable responses that did not align clearly with scoring rules were scored by the two coders by consensus.

Including all nine items, Cronbach's alpha was .665, and two items had low corrected item-total correlations (less than .2): Items 4 and 6a. CFA including all items had relatively good model fit, $X^2(27, N = 287) = 70.4$, $p < .001$, CFI = .90, RMSEA = .075. In line with item-total correlations, Items 4 and 6a had low factor loadings (less than .3). On the basis of these results, we decided not to include the scores for Items 4 and 6a in the final score for this scale. Including the final seven items, Cronbach's alpha was .679, and all corrected item-total correlations were above .2. The final CFA still had reasonably good model fit, $X^2(29, N = 287) = 95.1$, $p < .001$, CFI = .85, RMSEA = .089, and all factor loadings were above .4. The final resulting scale, using a sum score, had a normal distribution (range from 0 to 14, mean of 6.06, standard deviation of 3.39) with no indication of skewness or kurtosis (0.03 and -0.74, respectively).

Colors, Shapes, and Uses

For Colors, Shapes, and Uses (see Table D12), coders scored one open-ended item with possible scores ranging from 0 to 5: children's verbal definitions of what it means "to sort" (Item 2A). Interrater reliability for this item was high, with a weighted kappa of .94. All other items (1A, 1B, 1C, 1D, 2B, 3, and 4) were more concrete (involving either choosing particular objects, or one- to two-word responses) and were scored according to scoring rules using syntax. Any items with responses that did not conform clearly to scoring rules were scored by a coder. Twenty percent of items scored by syntax were also checked to ensure accurate scoring.

Including all eight items, Cronbach's alpha was .572. Although this did not indicate good internal consistency, all items had corrected item-total correlations above .2; in other words, no items stood out as not fitting with the scale. CFA, however, showed good model fit, $X^2(19, N = 275) = 32.2$, $p = .041$, CFI = .92, RMSEA = .047, and all items had good factor loadings, with one at .375 and all other items above .4. The final resulting scale had a normal distribution (range from 0 to 14, mean of 6.06, standard deviation of 3.39) with no indication of skewness or kurtosis (0.03 and -0.74, respectively). All items were retained for calculating the final score. The final resulting scale, using a sum score, had a normal distribution (range from 0 to 8, mean of 4.36, standard deviation of 1.36) with no indication of skewness or kurtosis (-0.34 and -0.10, respectively).

Parent Measures

Parent/Caregiver Pre- and PostSurveys

To develop the parent/caregiver pre- and postsurveys, the study team used some previously validated items from our recent national survey of parents about science and media use (Silander et al., 2018). In addition, the study team conducted cognitive interviews with five parents within the target sample (participants in the pilot study of the researcher-developed assessments) to refine the survey items.

Media Logs

The study team developed media logs based on media logs used previously during the Peg + Cat Content Study (Pasnik et al., 2015). The study team determined that an electronic log was more likely to yield a higher response rate and more accurate responses than a paper log returned at the end of the study. The log was distributed through the survey platform Qualtrics and texted to parents each Friday. The log asked parents to estimate how much time their child spent engaging with the *Cat in the Hat* digital resources, to report how they most often spent their time while their child engaged with the media, whether they used the hands-on activities, and whether their child was interested in the content.

Parent Interviews

The study team first developed a set of open-ended questions with the purpose of yielding rich responses about parent and child study experiences. Following interviews with parents at the first two of the five study sites (New York and Phoenix) that yielded insufficient details from parents, the study team revised the parent interview protocol to incorporate questions that were tailored to elicit more concrete examples and detailed responses from parents about both challenges and benefits of using the resources.

EDC Education Development Center

SRI Education™
A DIVISION OF SRI INTERNATIONAL