Do objects of different weights fall at the same time?  
Learning about gravity from picture books.

by

Purneshwarie Vaunam Venkadasalam

A thesis submitted in conformity with the requirements  
for the degree of Master of Arts  
Applied Psychology & Human Development  
University of Toronto

© Copyright by Purneshwarie Vaunam Venkadasalam 2016
Do objects of different weights fall at the same time? Learning about gravity from picture books.

Purneshwarie Vaunam Venkadasalam

Master of Arts

Applied Psychology & Human Development
University of Toronto

2016

Abstract

Picture books are a primary source of knowledge for children. However, young children are rarely exposed to informational books and spend more time reading fiction. We examined how children age 4 and 5 learn that objects with different weights fall at the same rate from picture books. A realistic fiction and an informational book with identical images were designed and matched in words and reading difficulty level. A pretest and posttest were administered to 120 children. Each test contained 4 pairs of objects identical in size, where two pairs of objects had the same weight and two pairs of objects had different weight. Although generally all children improved from pretest to posttest, 5-year-olds learnt from both types of book genres, while 4-year-olds learnt only from the realistic fictional book. Finally, children reported enjoying the realistic fictional and informational genres equally.
Acknowledgments

First I would like to thank my family for their support in helping me complete this project. To my sisters Valli and Mallar, thank you for all your honest advice and your proof reading. Secondly my lab mates, your advice has been invaluable. Thank you for always taking the time to have discussions with me and in helping me create the picture books for this study. I would also like to say a special thank you to Shelley He for her wonderful illustrations! Also, thank you to all the children and families that participated in this study. Lastly, I would like to thank my advisor Patricia Ganea, without you, this thesis would not have been possible.
# Table of Contents

**ACKNOWLEDGMENTS** ........................................................................................................... III

**TABLE OF CONTENTS** ............................................................................................................ IV

**LIST OF TABLES** ................................................................................................................ VI

**LIST OF FIGURES** ................................................................................................................ VII

**LIST OF APPENDICES** .......................................................................................................... VIII

### 1 REVIEW OF THE LITERATURE AND THEORETICAL BACKGROUND ................................. 1

1.1 **SCIENCE & LITERACY: TEACHING SCIENCE WITH BOOKS** ........................................ 1

1.2 **GENRE: NARRATIVE AND INFORMATIONAL BOOKS** .................................................... 2

   1.2.1 Conversations: Adult-Child Interactions & Peer-Talk .................................................... 3

   1.2.2 Memory & Comprehension ............................................................................................. 3

   1.2.3 Creating Books for Comparison ..................................................................................... 4

1.3 **INFORMATIONAL AND NARRATIVE: TWO PERSPECTIVES** ...................................... 5

   1.3.1 Narrative viewpoint ......................................................................................................... 5

   1.3.2 Informational Viewpoint ................................................................................................. 6

### 2 CURRENT STUDY .................................................................................................................. 7

2.1 **MISCONCEPTION ABOUT WEIGHT & GRAVITY** ............................................................ 8

2.2 **RATIONALE** ..................................................................................................................... 9

### 3 METHODS ............................................................................................................................ 10

3.1 **PARTICIPANTS** ............................................................................................................... 10

3.2 **MATERIALS** .................................................................................................................... 11

   3.2.1 Free Fall Stimuli ............................................................................................................. 11

   3.2.2 Picture Books ............................................................................................................... 11

### 4 PROCEDURE ......................................................................................................................... 11

4.1 **PRETEST PHASE** ............................................................................................................. 12

4.2 **PICTURE BOOK READING** ............................................................................................. 13

4.3 **COMPREHENSION QUESTIONS** ..................................................................................... 13

4.4 **POSTTEST** ....................................................................................................................... 14

4.5 **RECEPTIVE LANGUAGE ASSESSMENT** ....................................................................... 14

4.6 **CODING AND RELIABILITY** ......................................................................................... 14

### 5 RESULTS ................................................................................................................................ 15
5.1 LEARNING PHYSICAL SCIENCE FROM BOOKS ............................................................... 15
  5.1.1 Main Effects ........................................................................................................... 15
  5.1.2 Interactions .......................................................................................................... 16
5.2 RELATIVITY TO CHANCE ......................................................................................... 18
  5.2.1 Realistic Fiction Condition .................................................................................... 18
  5.2.2 Informational Condition ....................................................................................... 19
  5.2.3 Control Condition ............................................................................................... 19
5.3 BOOK PREFERENCES ............................................................................................... 20
  5.3.1 Open-Ended Comprehension Question .................................................................. 20
  5.3.2 Book Enjoyment Score ........................................................................................ 20
  5.3.3 Book Comprehension .......................................................................................... 21
  5.3.4 Reading Fluency ................................................................................................. 21
  5.3.5 Receptive language .............................................................................................. 21
6 DISCUSSION .................................................................................................................. 22
  6.1 LIMITATIONS .......................................................................................................... 23
  6.2 FUTURE STUDIES .................................................................................................... 24
  6.3 IMPLICATIONS ......................................................................................................... 25
7 CONCLUSION .................................................................................................................. 25
REFERENCES ..................................................................................................................... 27
TABLES .............................................................................................................................. 34
FIGURES ............................................................................................................................ 43
APPENDIX .......................................................................................................................... 44
List of Tables

1  TABLE 1: WEIGHT OF OBJECTS USED IN THE PRETEST AND POSTTEST .................................................. 34
2  TABLE 2: PERCENTAGE OF CHILDREN’S RESPONSES TO QUESTIONS ABOUT OBJECTS ............................... 35
3  TABLE 3: COMPREHENSION QUESTIONS FOR THE EXPERIMENTAL BOOKS ........................................ 36
4  TABLE 4: MEAN PRETEST AND POSTTEST SCORE BY BOOK CONDITION AND AGE .................................. 37
5  TABLE 5: PARTICIPANTS’ RESPONSE TO THE OPEN-ENDED QUESTIONS FOR THE EXPERIMENTAL BOOKS ... 38
6  TABLE 6: MEAN BOOK ENJOYMENT COMPOSITE SCORE BY BOOK CONDITION AND AGE ....................... 39
7  TABLE 7: MEAN BOOK COMPREHENSION SCORE BY BOOK CONDITION AND AGE .................................... 40
8  TABLE 8: MEAN READING FLUENCY SCORE BY BOOK CONDITION AND AGE ........................................ 41
9  TABLE 9: MEAN AGE-ADJUSTED RECEPTIVE LANGUAGE (TPVT) SCORE BY BOOK CONDITION AND AGE ...... 42
List of Figures

1  FIGURE 1. SAME WEIGHT OBJECTS AND DIFFERENT WEIGHT OBJECTS. ......................................................... 43
List of Appendices

1  REALISTIC FICTION AND INFORMATIONAL STORY SCRIPTS ................................................................. 44
Teaching science to children can be a difficult task for parents and educators alike. Today there is a heightened awareness of the importance in teaching Science, Technology, Engineering, and Mathematics (STEM) both in research and the educational system (DeJarnette, 2012). It is becoming increasingly important for children to be exposed to STEM topics earlier, to not only foster interest during childhood but also for later career trajectories (Mantzicopoulos & Patrick, 2011; Sinno, Schuette, & Killen, 2013). There are many ways to obtain children's interest and enrich their scientific knowledge, though books are commonly encountered as a teaching tool (Caswell & Duke, 1998). Children’s science books often have information that is not accessible from direct observation or through other forms of media (Pringle & Lamme, 2005). This study will examine how narrative and informational genres improve children’s abilities to learn a physics science concept.

1 Review of the Literature and Theoretical Background

This section will focus on research related to children’s literature in three main areas. First, the relationship between literacy and science will be examined. Next, research regarding book genres will be reviewed. Finally, two major theories dealing with the role of book genre in children’s science learning from picture books will be considered.

1.1 Science & Literacy: Teaching Science with Books

There is a long standing relationship between literacy and scientific knowledge, since “literary processes are the means by which science content is learned because content information is rooted in written and oral language” (Casteel & Isom, 1994, p. 540). Throughout recorded history, science knowledge has been immortalized in writing, which is then used to teach new pupils. Not only do books illustrate the theoretical aspects of science but books can also teach practical scientific ability. In fact children’s picture books can be used to create meaningful contexts that develop scientific skills such as making observations, predictions, classifications, measurements and even foster the development of communicating in scientific language (Monhardt & Monhardt, 2006). Empirical studies have supported this connection between science and literacy. Over the course of a school year, third grade children either participated in a literacy-science or literacy-only achievement group. The literacy-only classes used traditional books to teach children story retelling, story rewriting and comprehension skills, while the
literacy-science classes did the same using both science textbooks as well as traditional books. After the intervention, the literacy-science and literacy-only groups improved on all literacy measures and most science scores compared to the control group (Morrow et al., 1997). Surprisingly, children in both groups improved in some science measures, even though the literacy group was not superficially exposed to science material. Thereby, developing literacy skills indirectly improves the understanding of science. This demonstrates the importance that literacy plays in the role of science learning.

It is not surprising that studies have investigated the positive effect picture books can have on science learning. There is evidence that children can learn specific scientific ideas from picture books. A study with children aged between 5 and 8 used picture books to teach children about natural selection, a concept that is usually reserved for high school (Kelemen, Emmons, Seston Schillaci & Ganea, 2014). Young children's knowledge of natural selection increased from 18% to 87% after the storybook intervention. Another study demonstrated that even younger children can learn and transfer biological facts from picture books to real animals. Four-year-olds learned the concept of camouflage from picture books regardless of whether it was presented using factual or anthropomorphic language (Ganea, Ma, & DeLoache, 2011). These studies demonstrate that children are capable of learning science concepts from picture books. The next section examines how genre affects children's literacy development, with a specific focus on how different genres affect children's ability to learn science concepts.

1.2 Genre: Narrative and Informational Books

Genre plays an important role in determining how a book can be used in the classroom. Genre can be used to categorize a variety of different forms of media. With regard to literacy, genre is defined as the way texts are structured to serve different purposes in specific contexts (Kress, 1994). However, both narratives and informational books have been used to teach primary children science (Donovan & Smolkin, 2002). Therefore, the two types of genres that will be compared are narrative and informational books. Research that has investigated the effects of narrative and informational books on conversations as well as memory and comprehension will be reviewed below.
1.2.1 Conversations: Adult-Child Interactions & Peer-Talk

The breadth of research involving parent-child conversations in relation to language and literacy development is extensive (Ganea et al., 2011). However, not nearly as many studies have compared how information and narrative books promote discussions between children and their caregivers. Researchers have observed that 4-year-olds that engaged in joint reading with adults had considerably more interactions and extratextual talk when reading informational books (Anderson, Anderson, Lynch & Shapiro, 2004; Price, Kleeck, & Huberty, 2009), such that children used more diverse vocabulary and terminology, asked more questions that required an explanation and had more inferential reasoning interactions (Torr & Clugston, 1999; Price et al., 2009) compared to narratives. However, studies have also found that mothers were more likely to read a narrative book in its entirety compared to an informational book (Price et al., 2009). With toddlers, mothers also use more complex utterances during the narrative (Nyhout, & O’Neill, 2013). In addition to contrasting results within genres, previous studies have failed to find differences between the genres. Four-year-olds and parents produce more interactions while reading expository texts than narrative books but the quantity of high cognitive demand questions was similar between the two genres (Anderson, Anderson, Lynch, Shapiro & Eun Kim, 2012). Also, an informational storybook, which is a hybrid of a narrative and informational book, increased primary students’ discussions about the content and about extratextual topics, doubled the number of children’s speculations and also elicited more reciprocal replies when compared to informational books alone or storybooks (Leal, 1992). Overall, these results present conflicting views of how book genre affects conversations.

1.2.2 Memory & Comprehension

Research has investigated how children recall and understand narrative and expository texts. Some studies have found when children were asked to learn information about Earth from either an informational story (The Magic School Bus Inside Earth) or a science textbook, third graders retained more information (Leal, 1993a) and fifth graders recalled more information (Leal, 1993b) from the informational storybook than the science textbook. Fifth and seventh graders recalled relevant information better when it was present in either a narrative or an informational format, rather than from a hybrid text (Hidi, Baird, and Hildyard, 1982). However, this is not to say that children only learn from stories, since even after a single reading kindergarteners understood and could accurately paraphrase science-related informational text on life science,
earth, space and simple machines (Mantzicopoulos & Patrick, 2010). Furthermore, a correlative study found that after adding more expository books to children's reading collection, kindergarteners begin using more informational language (timeless verb constructions, generic noun constructions, repetition of theme, information-book-like beginnings etc.) when describing informational texts (Duke and Kay, 1998). Although it seems that young students may prefer story formats over expository books, other studies have shown children can also learn to utilize informational texts as well. This is illustrated by kindergarteners demonstrating the ability to recognize and comprehend narrative and informational texts after just three exposures to both genres (Pappas, 1993). With respect to primary grades, there were no differences in the factual information recalled about whales between the two genres immediately following the book reading or after a delay (Jetton, 1994). Yet again, these findings demonstrate that children can remember and comprehend information contained in both formats. Thus, there does not seem to be a clear advantage for recall and comprehension for any genre.

1.2.3 Creating Books for Comparison

It is important to note that very few of these studies controlled for topic or theme. Additionally, since all the studies used trade books there was considerable variability between the information being delivered between the two types of text. Only two studies have created pairs of books that were used to teach science concepts. The first study examined how genre affected fourth grade students’ fluency of reading, recall and comprehension of scientific ideas as well as text preference (Cervetti, Bravo, Hiebert, Pearson & Jaynes, 2009). Two pairs of informational and narrative books were created; one about the life cycle of a snail (life science) and one about how sand forms (Earth science). Students were similar in accuracy and rate of reading across the two types of genres and topics. Also children did not show a preference for either genre. However, children who read the informational text answered more comprehension questions correctly and recalled more key concepts than children who read the narrative book. In the second study, Arya and Maul (2012) examined the role of the scientific discovery narrative (SDN) in middle school science education. SDN presented science as a process of discovery from a first person viewpoint. The goal of this study was to ask whether SDN would have a positive impact on learning and whether there would be greater student text engagement. Two book pairs were created each consisting of an expository text and a SDN. The two topics covered were the Galilean telescope and radioactivity. Middle school students benefited from reading texts that
were written as discovery narratives. This was especially true for children from ethnically diverse and less socioeconomically advantaged backgrounds. The SND book supported the understanding and recall of conceptual information, but it is important to note that SDNs were longer than the expository texts. The authors suggested that the narrative increased interest and attention which allowed children to gain a deeper understanding and recall of the relevant information. Once more, the results of research to date support diverging perspectives.

1.3 Informational and Narrative: Two Perspectives

1.3.1 Narrative viewpoint

The overrepresentation of narratives can be easily explained since humans have particular sets of strategies that assist in remembering goal directed actions and casual events associated with individuals (Hidi et al., 1982). Some scholars consider that the best genre for 5- to 7-year olds is fiction since it connects to their emotions, is entertaining, expands their imaginations and facilitates perspective taking, all the while providing ways of exploring important issues and developing lasting reading habits (Keehn, Martinez & Teale, 2004). The authors stated that strong, problem-centered plot lines and stories featuring humorous situations are more likely to captivate children's attention because they will ponder what will occur next (Keehn et al., 2004). Children's deepest comprehension and most meaningful experiences arise from aesthetic responses from reliving stories (Shine & Roser, 1999). Leal (1992) suggests that the dual nature of informational storybooks helps to increase peer discourse. When scientific information is embedded in a narrative, it creates gaps in children’s conceptual mapping network. The author suggests that these gaps cause children to distinguish factual from fantasy information. Elementary students retain more scientific information from scientific stories because this genre captures their interest and attention, thereby allowing children to have a deeper understanding and recollection of the relevant information, which develops scientific thinking (Leal 1993b; Arya & Maul, 2012).

Students can be taught information through different types of books such as textbooks or trade books. An elementary teacher, Kaser (2002), used a variety of literary material including Anne Frank: The Diary of a Young Girl to pique interest in grade three, four and five students and to teach them about astronomy. The goal is to use literacy to create a starting point for curiosity, and support critical thinking and problem solving (Kaser, 2002). Trade books not specifically
designed to teach science, both fictional and nonfictional, can be successfully included in the process of teaching science units (Casteel & Isom, 1994; Kaser, 2002; Atkinson, Matusevich, & Huber, 2009). Thus, it is not impossible for children to gain scientific insight from sources that are not textbooks. In fact, there is a large acceptance of using both fiction and nonfiction trade books as a resource for teaching science. Trade books can be used when designing hands on activities, teaching critical thinking by questioning the author and evaluating the evidence (Rice, 2002). Overall, narrative competence is critical for early literacy as well as cognitive development (Paris & Paris, 2003) and teaching science should not be restricted to textbooks.

1.3.2 Informational Viewpoint

There is a common assumption that children gain the ability to understand and compose stories before the ability to understand expository texts (Pappas, 1993; Duke & Kays, 1998; Donovan & Smolkin, 2001; Duke & Tower, 2004). Even though there has been a large increase in the publication of nonfiction titles for children, only a small portion of these are available for children younger than the third grade (Duke & Tower, 2004). In addition, teachers have a negative perception of informational texts and consider them “unfun” (Donovan & Smolkin, 2002, p. 518). Teachers’ use of basal texts for elementary students relies predominantly on fiction (66% of basal texts), while informational literature is the second largest genre (20% of basal texts); not nearly as substantial (Moss & Newton, 2002). Preschool teachers only use 5% of expository texts during read-alouds (Yopp & Yopp, 2006). It is more difficult for teachers to find nonfiction texts appropriate for young readers since award winning nonfiction texts tend to be targeted at upper elementary students (Duke, & Tower, 2004). Teachers tend to use hybrid texts in teaching science instead of incorporating informational texts (Donovan & Smolkin, 2001).

The ability to use an informational text is a critical skill that remains underdeveloped in many children (Duke, 2000). There is a scarcity of informational texts in the collections of school libraries, displayed in the classroom or used in activities, especially in lower SES communities (Duke, 2000; Duke, Bennett-Armistead & Roberts, 2003). One problem created is that the reliance on fictional narrative can foster a narrow set of skills that are insufficient for the comprehension of non-narrative books (Mantzicopoulos & Patrick, 2011). This is evident in older students when they make the transition to more expository texts. Scholars have linked children’s limited comprehension of informational texts to patterns such as the “fourth grade
“slump” that develop especially in low SES children (Chall, Jacobs, & Baldwin, 1990). To overcome these shortcomings, scholars have advocated for the deployment of informational texts in primary grades (Duke, 2000).

2 Current Study

This study will investigate these two perspectives by examining whether kindergarteners learn a physical science concept from a narrative and informational book. Although a great deal of science content is published through children’s literature (Rice & Rainsford, 1996; Atkinson et al., 2009), adults should be careful when selecting books to specifically teach science concepts (Mayer, 1995; Atkinson et al., 2009) because learning science is also learning the language and terminology (Gee, 2004; Pappas, 2006). A study that reviewed children’s science books found that many contained misconceptions, anthropomorphism and inaccurate illustrations (Sackes, Trundle & Flevares, 2009). Anthropomorphic characters may have hindered second graders’ ability to distinguish fact and fiction in informational storybooks in comparison to third and fourth graders (Brabham, Boyd, & Edgington, 2000). Anthropomorphism has also impeded young children’s ability to learn about animals (Ganea, Canfield, Simons-Ghafari, & Chou, 2013). Critical characteristics for books selected for the purpose of teaching science are realistic illustrations and accurate representation of the concept (Pringle & Lamme, 2005). Books without these qualities actually increase the learning of science misconceptions (Mayer, 1995; Rice & Rainsford, 1996, Rice, 2002). Beneficial scientific picture books should create curiosity, leading to more questions being asked and answers being pursued (Pringle & Lamme, 2005). Thus, it is important to carefully select picture books when trying to teach children science concepts.

A realistic fiction and a nonfictional picture book were designed and used to compare how children learn a physical science idea from different genres. Fictional books are narratives that deal with events, places, and people that are not factual (Latrobe, Brodie & White, 2002) while nonfiction is a broad category of any literary material that presents facts (Duke, & Tower, 2004; Duke & Billman, 2009). The four main genres used to teach science are fictional stories, informational stories, informational books and hybrids (Donovan & Smolkin, 2002). The two genres utilized in this study were realistic fictional and informational books. Realistic fiction is an untrue occurrence that could actually happen (Latrobe et al., 2002). Informational texts convey generic information about the world, and are identifiable by the presentation style (labels,
captions, headings etc.) and repetition of a theme or topic (Latrobe et al., 2002; Duke, & Tower, 2004). More books are being published about life science concepts than physical science concepts (Sackes et al., 2009; Smolkin et al., 2009) and previous studies have taught children in the realm of biology (Ganea et al., 2011). However, the theme in these books dealt with the domain of physical science, as it is often neglected in primary children. The picture books in this study were designed to correct the misconception that exists around the relationship between weight, gravity and falling objects.

2.1 Misconception about Weight & Gravity

Before formal schooling, children develop prior conceptions about the world that are not always scientifically correct. Gravity is a force of attraction (Asghar & Libarkin, 2010). This abstract concept is not easily comprehended in its entirety, even among adults. Only 21% of first year geology students had the correct concept of gravity and even fewer could explain why gravity exists on earth; most relied on the function of gravity in their explanation (Asghar & Libarkin, 2010). Even science and physics third year student teachers had difficulty defining physical forces (such as inertial mass, gravitation mass, gravitational force, gravity and space) and had many misconceptions associated with these concepts (Gönen, 2008). Further, misconceptions about gravity have been found in middle and high school students (Galili, 1993; Palmer, 2001; Frappart, Raijmakers & Frède, 2014). For kindergarteners and first graders, falling is a natural motion. These children exhibit a falling error: they claim that a stone would fall towards a celestial body regardless of situational context because they conceive gravity as a property of the earth (Frappart, et al., 2014) not a force applicable to all masses. These predictions arise from children’s observations and experiences in the world. However, this assembled knowledge about earth and gravity is often fragmented, which is witnessed in the combination of accurate and inaccurate information children provide (Hannust, & Kikas, 2010).

Several studies investigated children’s naive theories of motion and found erroneous beliefs specifically regarding weight and free falling. Researchers have taught kindergarteners and first graders about the shape of the Earth, the relativity of direction (up and down) and to some degree gravity (Kikas et al., 2002). A misconception held by numerous individuals of a variety of ages is that heavier objects fall faster than lighter ones (Kavanagh, & Sneider, 2007; Hast, 2014). Not until 1687 when Isaac Newton published his three laws of motion could the notion that objects of
different weights will fall at the same rate be explained (Kavanagh & Sneider, 2007). Hast and Howe (2012) interviewed 144 children between the ages of 5 and 11 about their common-sense theories of motion and speed. In general, children associated faster motion with heaviness over lightness, roundness over other shapes, and smoothness over roughness. With respect to falling, children generally associated faster motion with heavier objects across all age levels. Mass accounted for almost all justifications for vertical motion. Alternatively, mass justified for a quarter of horizontal motion and even less for incline motion. In a further study with the same sample, Hast and Howe (2013) found that children to some degree understood that objects accelerate while falling. Children predicted that a heavier ball would accelerate when falling, which is consistent with the association between mass and falling motion found in the previous study.

2.2 Rationale

Trade books have traditionally been used to compare fictional and informational books and as a result these studies could not control for other factors such as text, number of pages, theme and content. In addition, very few studies have created similar texts while examining the effect genre has on learning science concepts (Donovan & Smolkin, 2002; Duke & Bilman, 2009). This has made it difficult to draw strong conclusions about the effect of genre without considering the variability between the trade books as a factor. The goal of this study is to examine whether book genre plays a role in how children learn physical science concepts from picture books. A realistic fiction and an informational book were designed to teach children that objects of different weights fall at the same rate. The books were comparable on several factors such as illustrations, readability, number of words and pages.

Since this research is novel and previous empirical evidence is inconsistent, the hypothesis for the current study is not explicitly apparent. One possibility is that realistic fiction books are more common and will promote learning, since humans process goal directed and causal events with particular strategies (Hidi et. al, 1982). Alternatively, the informational condition will outperform their counterparts because expository books will help in teaching scientific language to children (Pappas, 2006) and children do not have to keep track of the story while extracting information. Finally, there could be no differences in learning between the two genres. There may be preferences toward one genre, since older children and adolescent boys generally prefer
nonfiction and physical science topics, while girls enjoy fictional texts and biological related themes (Mantzicopoulos & Patrick, 2011). The first outcome could be that males prefer informational texts over realistic fiction and girls have a higher preference rating for realistic fiction over the informational book. The alternative hypothesis is that males will prefer the realistic fiction, while females will enjoy the informational book. Nevertheless, there may be no gender differences toward a genre.

3 Methods

3.1 Participants

One hundred forty four children ages 4 and 5 were tested in three conditions; 40 children in the realistic fiction ($M = 5.03, SD = .59; 21 females and 19 males), 40 in the informational ($M = 5.05, SD = .59; 19 females and 21 males) and 40 in the control ($M = 5.05, SD = .55; 19 females and 21 males). Twenty five additional children were excluded because they passed the pretest ($N = 8$), failed the comprehension questions in the experimental conditions ($N = 8$), had a receptive language score that was more than two standard deviations below the mean ($N = 4$), received feedback by dropping the objects ($N = 2$) or due to experimenter error ($N = 3$).

Participants were recruited from either a database of families that have expressed interest in being contacted to take part in research at the university or recruited at the Ontario Science Center. The majority of children were White (54%) but the remaining children were South Asian (9%), Chinese (7%), Black (2%), Filipino (1%), Latin American (1%), Arabian (1%), Southeast Asian (1%) or from mixed ethnicities (18%). An additional 8% of families did not disclose their background. With respect to parental education, 13% of the sample did not report level of schooling completed. Among parents who reported their educational background almost all earned at least a high school degree (99%). A preponderance of parents earned a bachelor's degree (38%) or a master's degree (31%). The remaining parents either obtained a diploma (14%), doctoral degree (9%) or just a high school degree (6%).
3.2 Materials

3.2.1 Free Fall Stimuli

Eight pairs of objects were created (Figure 1). Four pairs had the same weight while four pairs had different weights. The weight was distinctly different since one object was left empty and the corresponding object was filled with glue and small pebbles (see Table 1 for exact weight). The objects were designed so that half of the same weight objects had two pairs that looked identical and two that looked different. Similarly, there were two identical and non-identical different weight pairs. The objects were organized into two groups so children received one pair of each type in the pretest and posttest: identical same weight, non-identical different weight, non-identical same weight and identical different weight. The order was counterbalanced so half the children were tested with objects from Group A then Group B in the pretest and posttest respectively and vice versa.

3.2.2 Picture Books

Two books (23 cm x 29 cm) with identical illustrations were created to teach participants about weight, falling objects and gravity. The realistic fiction and informational books were similar in number of pages (14 each), word count (419 and 407), Flesch Kincaid Reading Ease (98 and 92) and Automated Readability Index (2.4 and 3.9). Using a readability tool, both books were scored very easy to read and were suitable for children eight to nine years of age to read independently. The experimental books were designed to be analogous in every aspect with the exception of how the information was delivered (Appendix). The control book was based on the children's book Plants Feed Me by Lizzy Rockwell. The plant book (23 cm x 29 cm) was adapted to match the experimental books as it was also scored very easy to read and was appropriate for the same age range. The book matched in length (14 pages), word count (416), Flesch Kincaid Reading Ease (93) and Automated Readability Index (3.8). A PhD engineering graduate student read the books for accuracy.

4 Procedure

The four phases in this study were a pretest phase, a picture book reading followed by comprehension questions and lastly a posttest phase. The entire session lasted approximately 20 to 30 minutes and was conducted by the same experimenter. Children were either tested in the
lab at the university or at the Ontario Science Center. Children were randomly assigned to one of three book conditions; the realistic fiction, informational or control. Participants began by either colouring pictures or completing a short puzzle with the experimenter for a few minutes in order to make the children relax and feel at ease with the experimenter. Once children appeared comfortable the experimenter proceeded to the experiment.

The experimenter stated she had some objects for the child to look at and would ask some questions about the objects. First, children were told to hold two objects with different weights, one in each hand. The experimenter then asked children 'Do you think these objects have the same weight or different weight?'. Children who answered incorrectly were asked which object they thought was heavier and which one was lighter. The experimenter then went on to explain that if one object is heavier than another then it means that the objects have different weights. Children who answered correctly proceeded directly to the next pair of objects which were equal in weight. The same procedure was repeated using a same weight pair. This first task served as an assessment of the children’s knowledge of weight. Children who did not correctly answer one or both of the questions were asked the questions again. Once children demonstrated an understanding of weight they continued to the pretest.

4.1 Pretest Phase

Participants were given an identical pair of objects to inspect. To help children focus their attention on the aspect of size and weight, children were explicitly asked about these features: 'Do you think these objects have the same size?' and 'Do they have the same weight?'. Depending on the response 'Which one is heavier?' was asked. The order of these questions was counter balanced across the test phases and conditions and no feedback was given to the children regardless of accuracy. Children were proficient at correctly assessing the object's size, weight and identifying the heavier object with an average accuracy rate of 88.3%, 97.1% and 98.8% respectively (See Table 2 for individual objects). Finally, children were asked the test question which was adapted from Hast (2012): 'If you hold the objects out like this and let them drop at the same time, do you think one of the two will fall faster or do you think they will both fall at the same time?'. The sequence of the question was alternated so that half the time children heard “both fall at the same time” first, followed by “one of the two will fall faster”. Children's responses were recorded and the experimenter replied with a neutral response, such "okay". The
process was repeated for the remaining three pairs of objects (non-identical different weight, non-identical same weight and identical different weight).

4.2 Picture Book Reading

Children were read either a realistic fiction, informational or control book. All books were introduced uniformly; “Next we are going to read a book called…”. If children asked questions during the book reading, the experimenter replied neutrally and directed their attention back to the book. Each book was read twice to all children except one. This child declined to reading the book again, but answered four out of five of the comprehension questions correctly and was included in the analysis.

4.3 Comprehension Questions

Following the book reading session, five comprehension questions were asked with regard to the experimental books (see Table 3). The comprehension questions verify that children both paid attention and understood the material that was read to them. Eight children were excluded because they failed to answer three out of five questions correctly. An open-ended question was asked in the experimental conditions as well as the control: ‘Can you tell me something that happened in the book?’. If children failed to respond they were prompted with ‘Can you think of one thing that you remember from the book?’ followed by ‘What was the book about? Did you learn anything from the book?’. Over half of the participants provided a reply to the first question (58%), almost a third to the second query (32%) and the remaining to the third (11%). This was distributed consistently across all three conditions. The open-ended question was initially designed to examine if children would spontaneously mention the information about how weight is not a factor of falling objects. For the control group, the open-ended question served as a measure of attention. Further, book preference was also assessed. A composite score was produced from how children rated how much they enjoyed reading the book and the likelihood of reading the book again if they had it at home. Participants chose their answer from one of the following: not at all, not really, can’t decide, a little, or a lot. The posttest followed this phase next.
4.4 Posttest

Children completed the same task as the pretest with the second group of objects in the same order. This permitted a direct evaluation of learning as a result of the picture book reading.

4.5 Receptive Language Assessment

Children's receptive language score was also measured using the NIH Toolbox Picture Vocabulary Test. This is a computer adaptive task where children see four pictures on the screen and have to select the picture that best matches the word they hear. The age adjusted score was generated by the program and used to assess children's receptive language. Three children were reluctant to complete the TPVT though they demonstrated adequate understanding of English through answers to the comprehension questions and parental report of language. The task was used to ensure that children understood the instruction of the task and even more crucially the picture book. As a result, four children were excluded because their age adjusted score was more than two standard deviations below the mean.

4.6 Coding and Reliability

Children's answers to the test question were recorded by the experimenter during the experiment. The correct answer is that both objects would fall at the same time. An undergraduate research assistant coded 96% of the children's responses from the video (five videos cut out or were not filmed successfully). The interrater reliability between the live coder and research assistant was found to be Kappa = 0.96 with a 99% agreement between the raters. The disagreement was resolved by discussion.

The open-ended question was also coded by two research assistants using a 4-point scale. The interrater reliability was found to be Kappa = 0.81 with an 89% agreement between the raters. Again disagreements were resolved by discussion.

1. Conceptual understanding: Children's responses that explained that objects fell at the same time and indicated that mass was not a factor were categorized as having "conceptual understanding". For example "The buckets got down at the same time even though they were different weights".
2. Objects Fall down together: These included responses that said that objects would fall at the same time without making a reference to mass. For example "They dropped them and all landed at the same time" or "They fall down at the same time".

3. Book related: These answers were ones that referred to events from the book (e.g. "The ice cream fell down") or said the title of the book (e.g. "All things fall down").

4. Other: Any other response not categorized as above was coded as other, including "I don't know" or no response.

5 Results
The first analysis conducted used an ANOVA to investigate whether children learned that objects fall at the same time regardless of weight from the picture book conditions and considered age and gender as factors. Next, children’s responses were examined in comparison to chance. Finally, an analysis considered if there was a preference for a book genre.

5.1 Learning Physical Science from Books
A mixed-design analysis of variance (ANOVA), with test phase (pretest, posttest) as the within-subjects factor and picture book condition (realistic fiction, informational, control), gender (female, male) and age (4, 5) as the between-subjects factors. The homogeneity of variances was assessed by the Levene's test of homogeneity of variance for the pretest ($p = .06$) and posttest ($p = .08$) scores. The homogeneity of covariances was assessed by Box's test of equality of covariance matrices ($p = .07$). The mixed ANOVA revealed a main effect of test phase. There was a statistically significant difference in mean scores from pretest to posttest, $F(1, 108) = 64.57, p < .001$, partial $\eta^2 = .37$. Children's scores in the posttest ($M = 2.42$, $SD = 0.10$) showed a mean increase of 0.86, 95% CI [0.64, 1.07] from pretest ($M = 1.57$, $SD = 0.07$). The mean score of responses given by the children in the two age groups as a function of test phase and book condition is summarized in Table 4.

5.1.1 Main Effects
The main effect of book condition showed that there was a statistically significant difference in test phase, $F(2, 108) = 18.87, p < .001$, partial $\eta^2 = .26$. The control condition test phase scores ($M = 1.43$, $SD = 0.12$) were lower than the realistic fiction ($M = 2.41$, $SD = 0.12$) and
informational ($M = 2.14$, $SD = 0.12$) conditions. Tukey post hoc analysis revealed that the test phase mean increased from control to realistic fiction (0.99, 95% CI [0.58, 1.39]) and was statistically significant ($p < .001$). There was an increase from control to informational conditions (0.73, 95% CI [0.32, 1.13], $p < .001$), but the test phase differences between realistic fiction and informational (0.26, 95% CI [-0.14, 0.67]) were not statistically significant ($p = .348$).

There was no main effect of age, $F(1, 108) = 1.53$, $p = .22$, partial $\eta^2 = .01$ or gender, $F(1, 108) = 2.02$, $p = .16$, partial $\eta^2 = .02$. Finally, there was no interaction between condition and age, $F(2, 108) = 0.86$, $p = .43$, partial $\eta^2 = .02$, condition and gender, $F(2, 108) = 0.40$, $p = .67$, partial $\eta^2 = .007$, age or gender, $F(1, 108) = 0.40$, $p = .53$, partial $\eta^2 = .004$. The three way interaction between condition, age and gender was also not significant, $F(2, 108) = 0.03$, $p = .97$, partial $\eta^2 = .00$.

5.1.2 Interactions

The mixed ANOVA displayed a statistically significant interaction between test phase and picture book condition, $F(2, 108) = 3.93$, $p < .02$, partial $\eta^2 = .07$. There was also a significant interaction between test phase and age, $F(1, 108) = 5.92$, $p < .02$, partial $\eta^2 = .05$, and between test phase and gender, $F(1, 108) = 4.38$, $p < .04$, partial $\eta^2 = .04$. There were no significant three way interactions: test phase, book condition, age ($p = .78$), test phase, book condition, gender ($p = .09$), test phase, age, gender ($p = .19$). The four way interaction between test phase, book condition, age and gender was also not significant ($p = .38$).

5.1.2.1 Picture Book Condition Interactions

To examine the interaction between test phase and picture book conditions, two separate analyses of variances were conducted for the pretest and posttest respectively. The first ANOVA revealed a statistically significant difference in pretest scores between the picture book condition, $F(2, 117) = 8.50$, $p < .001$, partial $\eta^2 = .13$. Post hoc comparisons using the Bonferroni test revealed that children in the control condition scored significantly less on the pretest than children in the realistic fiction ($M = 0.68$, $p < .001$, 95% CI = -1.08, -0.27) and informational books ($M = 0.43$, $p = .04$, 95% CI = -0.83, -0.02). There was no difference between the realistic fiction and informational books ($M = 0.25$, $p = .40$, 95% CI = -0.15, 0.65). The second ANOVA also found a statistically significant difference in posttest scores between book conditions, $F(2, 117) = 14.87$, $p < .001$, partial $\eta^2 = .20$. As before, the Bonferroni post hoc test revealed that
children in the control condition scored significantly less on the posttest than children in the realistic fiction ($M = 1.30$, $p < .001$, 95% CI = -1.80, -0.80) and informational books ($M = 1.03$, $p < .001$, 95% CI = -1.52, -0.53). There was no difference between the realistic fiction and informational book ($M = 0.27$, $p = .28$, 95% CI = -0.77, 0.22). A repeated measure ANOVA for test phase was conducted for each book condition. There was a statistically significant effect of learning for the realistic fiction condition, $F(1, 39) = 27.00$, $p < .001$, partial $\eta^2 = .41$, informational condition, $F(1, 39) = 21.53$, $p < .001$, partial $\eta^2 = .36$, and control, $F(1, 39) = 12.20$, $p = .001$, partial $\eta^2 = .24$. All three conditions increased mean scores from the pretest to posttest. However, the realistic fiction ($M = 2.41$, 95% CI = 2.18, 2.64) and informational conditions ($M = 2.15$, 95% CI = 1.95, 2.40) improved more than the control ($M = 1.43$, 95% CI = 1.16, 1.70).

5.1.2.2 Age Group Interactions

An independent-samples $t$-test was conducted to examine the interaction between test phase and age. There was no difference in pretest scores between 4-year-olds ($M = 1.61$, $SD = .87$) and 5-year-olds ($M = 1.52$, $SD = .70$), $t(1, 118) = 0.60$, $p = .55$, $d = .11$, 95% CI [-0.20 to 0.37]. Conversely, the 4-year-olds ($M = 2.20$, $SD = 1.20$) posttest scores were lower than the 5-year-olds ($M = 2.64$, $SD = 1.27$), a marginally significant difference, $M = -0.44$, 95% CI [-0.88, 0.01], $t(118) = -1.94$, $p = .055$, $d = .36$. A paired samples $t$-test indicated that there was a statistically significant increase in scores from pretest to posttest for 4-year-olds, $t(58) = 4.33$, $p < .001$, $d = .56$ and 5-year-olds, $t(60) = 6.42$, $p < .001$, $d = .82$. Both ages increased mean scores from the pretest to posttest, but 5-year-olds ($M = 1.12$, 95% CI = 0.77, 1.46) had a larger increase in mean scores compared to 4-year-olds ($M = .59$, 95% CI = 0.32, 0.87).

5.1.2.3 Gender Interactions

An independent-samples $t$-test was conducted to examine the interaction between test phase and gender. There was no difference in pretest scores between females ($M = 1.56$, $SD = .79$) and males ($M = 1.57$, $SD = .78$), $t(1, 118) = -0.10$, $p = .92$, $d = 0.01$, 95% CI [-0.30 to 0.27]. However, the females ($M = 2.66$, $SD = 1.25$) posttest scores were higher than males ($M = 2.22$, $SD = 1.21$), a marginally significant difference, $M = 0.41$, 95% CI [0.19, 0.91], $t(118) = 2.07$, $p = .04$, $d = 0.37$. A paired samples $t$-test indicated that there was a statistically significant increase in scores from pretest to posttest for females, $t(58) = 6.74$, $p < .001$, $d = .88$ and males,
Both ages increased mean scores from the pretest to posttest, but females ($M = 1.10$, $95\%$ CI $= 0.78$, 1.43) had a larger increase in mean scores compared to males ($M = .62$, $95\%$ CI $= 0.32$, 0.93).

5.2 Relativity to Chance

Since the control group was also demonstrating significant improvements from pretest to posttest, a one-sample $t$-test was conducted to determine if the pretest and posttest score was different than chance for each condition and age group. There were four test questions in each test phase, therefore chance is equal to two for all tests. The result for each condition and age are reported below.

5.2.1 Realistic Fiction Condition

5.2.1.1 Four-year-olds

The pretest score for 4-year-olds ($M = 2.10$, $SD = 0.71$) was not significantly different than chance, the mean difference was $0.10$, $t(19) = .62$, $p = .54$, $d = .14$, $95\%$ CI [-0.24, 0.44]. Conversely, the posttest score ($M = 2.80$, $SD = 1.11$) was greater than chance, a statistically significant difference of $0.80$ ($95\%$ CI, $0.28$ to $1.32$), $t(19) = 3.24$, $p = .004$, $d = .72$. This suggests that 4-year-olds did learn from the realistic fiction book, since they improved from chance levels at the pretest to above chance levels in the posttest.

5.2.1.2 Five-year-olds

The pretest score for 5-year-olds ($M = 1.65$, $SD = 0.67$) was lower than chance, a statistically significant difference of $-0.35$ ($95\%$ CI, $-0.66$ to $-0.04$), $t(19) = -2.33$, $p = .03$, $d = .52$. Additionally, the posttest score ($M = 3.10$, $SD = 1.25$) was greater than chance, a statistically significant difference of $1.10$ ($95\%$ CI, $0.51$ 8 to $1.69$), $t(19) = 3.93$, $p = .001$, $d = .88$. This implies that 5-year-olds learned from the realistic fiction book as well. Five-year-olds scores improved from below chance to above chance from the pretest to the posttest.
5.2.2 Informational Condition

5.2.2.1 Four-year-olds

The pretest score for 4-year-olds ($M = 1.58$, $SD = 0.60$) was lower than chance, a statistically significant difference of -0.42 (95% CI, -0.71 to -0.13), $t(18) = -3.02$, $p = .007$, $d = .69$. However, the posttest score ($M = 2.42$, $SD = 1.07$) was not significantly different than chance, the mean difference was 0.42, $t(18) = 1.71$, $p = .10$, $d = .39$, 95% CI [0.09, 0.94]. These results contrast the realistic fiction, such that 4-year-olds did not learn from the informational picture book because the posttest score was not significantly different than chance.

5.2.2.2 Five-year-olds

The pretest score for 5-year-olds ($M = 1.67$, $SD = .66$) was lower than chance, a statistically significant difference of -0.33 (95% CI, -0.63 to -0.03), $t(20) = -2.32$, $p = .03$, $d = .51$. The posttest score was ($M = 2.90$, $SD = 1.26$) was greater than chance, a statistically significant difference of 0.91 (95% CI, 0.33 to 1.48), $t(20) = 3.29$, $p = .004$, $d = .72$. These outcomes are strikingly similar to 5-year-olds in the realistic fiction group, their scores improved from below chance in the pretest to above chance in the posttest. This suggests that 5-year-olds learned from both types of books.

5.2.3 Control Condition

5.2.3.1 Four-year-olds

The pretest score for 4-year-olds ($M = 1.15$, $SD = .99$) was lower than chance, a statistically significant difference of -0.85 (95% CI, -1.31 to -0.39), $t(19) = -3.85$, $p = .001$, $d = .86$. However, the posttest score ($M = 1.40$, $SD = 1.00$) was also lower than chance, a statistically significant difference of -0.60 (95% CI, -1.07 to -0.13), $t(19) = -2.70$, $p = .01$, $d = .60$. Although there was an improvement in mean score after the book reading, it does not indicate learning since the pretest and posttest scores were both significantly below chance.

5.2.3.2 Five-year-olds

The pretest score for 5-year-olds ($M = 1.25$, $SD = .72$) was lower than chance, a statistically significant difference of -0.75 (95% CI, -1.09 to -.41), $t(19) = -4.68$, $p < .001$, $d = 1.05$. However, the posttest score ($M = 1.90$, $SD = .97$) was not significantly different than chance, the mean
difference was 0.10, \( t(19) = -0.46, p = .65, d = .10, 95\% \text{ CI } [-0.55, 0.35] \). These findings imply that the 5-year-olds did not learn from the control books, since the posttest scores were not significantly different than chance.

### 5.3 Book Preferences

#### 5.3.1 Open-Ended Comprehension Question

Children in the experimental conditions were asked to recall something that happened in the book to explore whether children would explicitly provide details about the concept being taught (see Table 5). A substantial amount of replies was book related (61.3%). Very few children actually provided information regarding falling objects (17.5%). Only 10% of children explicitly explained the relationship between gravity, falling objects and weight (10%). A chi-square linear-by-linear association test was conducted between picture book condition (realistic fiction or informational) and the responses to the open-ended question (ordinal data). There was a statistically significant association between experimental book condition and children's responses, \( \chi^2(1) = 4.490, p = .04 \). A Spearman's rank-order correlation was conducted to assess the relationship between children’s categorized responses and the experimental book conditions. There was a weak correlation, \( r_s(78) = .23, p = .04 \). The realistic fiction condition had more answers categorized as level one, while the informational group had more answers categorized as level four.

#### 5.3.2 Book Enjoyment Score

Children's reported book enjoyment mean composite score are in Table 6 for each age and condition. A Spearman's rank-order correlation was conducted to assess the relationship between children's reported book enjoyment composite score and the type of picture book. There was no correlation between the book enjoyment score and the type of picture book children were read, \( r_s(118) = .18, p = .05 \). Similarly, a Spearman's rank-order correlation was used to see if there was a relationship between children's reported book enjoyment composite score and the mean posttest score. No correlation was found for posttest score and book enjoyment score, \( r_s(118) = .18, p = .05 \). Children's book enjoyment was not related to the type of book they read (realistic fiction, informational or control) or learning.
With respect to the experimental books (realistic fiction and informational) there was no correlation between the book enjoyment scores, $r_s(78) = .006, p = .96$. Children enjoyed both the realistic and informational book equally. Furthermore, no correlation was found for the posttest score and the experimental books, $r_s(78) = .12, p = .28$. Children enjoyment of the book was not associated with children's learning outcomes.

To investigate if there were any gender differences in book enjoyment, a Spearman's rank-order test was done and no correlation was found, $r_s(118) = -.05, p = .61$. Gender was not related to book enjoyment score for any of the three book conditions. Specifically for the experimental books, no correlation was found between gender and book enjoyment either, $r_s(78) = -.07, p = .53$. Males and females enjoyed both the realistic and informational books equally.

### 5.3.3 Book Comprehension

Recall that the book comprehension measure was only employed in the experimental book conditions. A Spearman's rank-order found a weak correlation between children's experimental book comprehension score and the posttest score, $r_s(78) = .26, p = .02$. As expected, children who scored higher on the book comprehension questions were more likely to have a higher posttest score. The children's reported book comprehension scores are in Table 7 for the realistic fiction and informational conditions as well as age.

### 5.3.4 Reading Fluency

A Spearman's rank-order found a weak correlation between children who read fluently and the posttest score, $r_s(110) = .25, p = .007$. This relationship was stronger when just considering the experimental book. A Spearman's rank-order found a correlation between children in the experimental groups who read fluently and the posttest score, $r_s(70) = .35, p = .002$. Parents who reported that their children could read fluently were more likely to have a higher posttest score than children whose parents reported they could not. For the average score for reading fluency for each condition and age group, see Table 8.

### 5.3.5 Receptive language

Children's age adjusted receptive language (TPVT) scores are in Table 9 for each age and condition. Finally, a Spearman's rank-order correlation was used to see if there was a relationship
between children’s receptive language score and the mean posttest score. No correlation was found, \( r_s(118) = -0.05, p = .61 \). Children’s receptive language was not associated with children's learning outcomes. Intriguingly, a Spearman's rank-order found a weak correlation between children's receptive language score and experimental group books comprehension score, \( r_s(76) = .24, p = .04 \). Thereby, children with higher receptive language were more likely to have a higher book comprehension score.

6 Discussion

Generally, trade books have been used to compare how children learn science concepts from different genres without controlling for other factors. Children were exposed to the concept in either a realistic fiction or an informational picture book that were designed to be as similar as possible. The objective of this study was to examine whether children would correct a common misconception that heavier objects fall at a faster rate after a picture book intervention. The results confirmed that there is a complex relationship among science learning outcomes and genres since no single genre was superior (Cervetti et al., 2009). The findings revealed that children learnt from both types of genres, but 5-year-olds learnt significantly more than 4-year-olds. Additionally, females learnt more than males as indicated in the difference between their posttest scores.

Findings from our chance analysis are inconsistent with the results from the ANOVA. Five-year-olds learnt that objects fall at the same rate from both types of books. In contrast, 4-year-olds appeared to learn from the realistic fiction book, however, their posttest score in the informational condition was not significantly different than chance. Therefore, the pretest-posttest difference is significant but it cannot be inferred that 4-year-olds systematically learn the concept as their posttest scores are at chance. Therefore, the improvement seen in the informational condition may not be attributed to learning. Similarly, the control group improvement from pretest to posttest was also not greater than chance.

There are two possible factors that can alternatively explain the gains in learning from pretest to posttest. The first possibility is that there might be a task effect. Children might be learning something from the falling task, which may account for a portion of test score improvement. However, since all children received the objects in the same order, the task effect is expected to be a factor across all conditions. Therefore, the increase in score from pretest and posttest is
likely not a task effect. Secondly, the picture book reading may be causing the improvement. The control group is to account for learning that happens outside of the experimental picture books. It was predicted that there should be no improvement in the control condition in comparison to the experimental condition. However, the control condition had a significant gain increase, yet the posttest score was not different than chance. This variation can be attributed to the significantly lower pretest score in the control group compared to the experimental conditions. Children were randomly assigned to conditions and therefore, the groups might not have been created equally and other factors may have contributed to the gain. As a result, the increase in the control was not greater than chance, whereas the 5-year-olds improved from below chance in the pretest phase to above chance in the posttest from both genres. This increase was seen in the realistic fiction for 4-year-olds, however, the informational group was not above chance. To summarize, children might be learning from the free fall task, although not enough to significantly improve their score above chance and the control group started at a lower starting point compared to the experimental picture book conditions. Therefore, 5-year-olds are learning from both types of books genres, while 4-year-olds seem to show significant learning from the realistic fiction.

The subsequent analysis compared book preferences. The null hypothesis is accepted as there were no book preference between the realistic fiction and informational book. This is inconsistent with previous research with older children since a preference was found for fiction (Moss & McDonald, 2004) and nonfiction (Caswell & Duke, 1998; Mohr, 2006). However, some studies have found no preference in fourth graders (Cervetti et al., 2009). These findings suggest that a genre may not be preferred over another, instead it is more important for reading material to appeal to students, especially those who struggle and abhor reading (Worthy, 1996).

6.1 Limitations

There are some drawbacks to this study. The first is that children could be guessing the answers when asked to predict whether the objects will fall down together or if one will fall faster than the other. However, children were not asked to explain their predictions. The open-ended question was initially designed to see if children could explain the concept of gravity from the picture books. Children's responses were fairly equal in the book related category and objects fall down category, while very few children verbalized the concept of free falling and gravity. Furthermore, anecdotally some children said the heavier shoe filled with sand begin losing some
of its substance while it was descending. Although, only a minority of children raised this concern, not all children may have vocalized this observation and we cannot predict the effects it has on the results.

Finally, not all participants may have had this misconception, since they were never explicitly asked. Children were organized into having the misconceptions or not, based on the pretest and posttest scores for the same and different weight objects. Not all children exhibited the misconception that heavier objects fall faster; 77.0% of 5-year-olds and 66.4% of 4-year-olds. Another common misconception younger children may have prior to demonstrating that heavier objects fall faster, is the justification that things fall when they are not supported (Kavanagh & Sneider, 2007). Bar and colleagues (1994) have found that some 4 and 5-year-old children do not consider the object's weight as a relevant variable when they predict what will happen when two objects similar in size are dropped together. These children predict that objects of different weight will reach the ground at the same time. It is between age 5 and seven that 50% of their sample used weight as a factor of falling. So it is possible within this sample that some children did not have this misconception, thus did not consider how the mass of the object will affect the rate of falling. Hast and Howe (2012) found that younger children relied more on size to justify an object's motion and did not consistently use mass the same way that older children did. This suggests perhaps that younger children begin to use size as a proxy for mass. If children do not associate faster motions with heavier objects but refer to size, they might reason faster motion is associated with bigger objects. In this study, all objects had the same size; therefore children could not use this feature to make a differentiation between how the objects fall.

6.2 Future Studies

Future studies could establish the misconception the child has prior to implementing an intervention, for example by asking ‘Why would a stone thrown up fall down again?’ (Kikas et al., 2002). Further, when children make their prediction, they could be asked to explain their choices and explicitly asked to explain the concept learned in the books. This would be an explicit measure of learning from the pretest to the posttest. Interventions could also include adult scaffolding. For younger children, discussing the content of the informational books may be effective for conveying science and social studies content. (Brabham, Boyd, & Edgington, 2000). Furthermore, there is a preference to using hybrid books in science especially in primary
classrooms, such as *The Magic School Bus* (Pappas, 2006), studies could investigate how children learn from hybrid fiction and informational books.

### 6.3 Implications

Engagement with science informational texts has the potential to fuel children’s long-term interest in science as well as educate them with scientific language (Mantzicopoulos & Patrick, 2011). Teachers think informational books are not fun (Donovan & Smolkin, 2002) in comparison to fictional books. Additionally, teachers tend to use narrative or hybrid books to teach science rather than incorporate informational texts reflecting scientific language (Donovan & Smolkin, 2001). However, it is imperative for teachers to provide a balance in the materials selected for use in the classroom. As of yet our educational system has not succeeded in bringing children adequate exposure to different genres (Duke, 2003; Duke & Tower, 2004). Teachers should be cautious when selecting science trade books (Smolkin, McTigue, Donovan & Coleman, 2009) and informational texts (Atkinson et al., 2009) for primary classrooms. The primary criteria for good informational texts for young readers is that the book is interesting but conveys at least some information children do not already know (Duke, & Tower, 2004). Other factors to consider include comprehensiveness (Smolkin et al., 2009) accuracy, appeal, and complexity (Tyson, 1999;Donovan & Smolkin, 2002; Duke, & Tower, 2004). The role of the teacher is essential in not only text breadth and evaluating appropriateness but teachers can also highlight the differences between genres so children can learn to use both types effectively (Duke, 2000; Calo, 2011).

### 7 Conclusion

Generally, very few physical science books are written for primary grades (Sackes et al., 2009; Smolkin et al., 2009). This study provided children with the opportunity to learn about a physical science concept through picture books, since children integrate the information contained in the pictures and texts with their prior knowledge (Paris & Paris, 2003), however this study demonstrated that acquiring knowledge through text is a complex process (Cervetti et al., 2009). Five-year-old children gained knowledge of gravity and improved from both the realistic fictional and informational genres. Although, 4-year-olds improved after reading both books, they only improved to levels greater than chance after reading the realistic fiction book. Furthermore, children had no preference of book genre. These results suggest that children are
capable of learning science from both types of books, but the realistic fiction may provide a superior benefit for children younger than 5 years of age. Picture books are a tool that can be utilized by teachers to fuel the interest of children in the area of science and eventually the pursuit of STEM careers.
References


## Tables

### 1 Table 1: Weight of Objects Used in the Pretest and Posttest

<table>
<thead>
<tr>
<th>Object Pairs</th>
<th>Group A Weight (grams)</th>
<th>Group B Weight (grams)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identical Same Weight</td>
<td>0.5</td>
<td>1.2</td>
</tr>
<tr>
<td>Non-identical Different Weight</td>
<td>1.2</td>
<td>5.5</td>
</tr>
<tr>
<td>Non-identical Same Weight</td>
<td>0.8</td>
<td>1.2</td>
</tr>
<tr>
<td>Identical Different Weight</td>
<td>1.8</td>
<td>9.4</td>
</tr>
</tbody>
</table>

*Notes.* Group A consisted of identical same weight small paint rollers (approximately 3.50 cm in diameter with a length of 11.4 cm), different weight paint rollers (approximately 3.60 cm in diameter with a length of 16.7 cm), same weight container (approximately 7.00 cm length × 7.00 cm width × 4.50 cm height) and different weight red egg boilers (approximately 6.30 cm in diameter with a length of 10.1 cm). Group B comprised of identical same weight balls (approximately 6.00 cm in diameter), different weight stacked containers (approximately 6.00 cm length × 6.00 cm width × 9.00 cm height), same weight hour glass shaped containers (approximately 6.50 cm in diameter with a height of 11.0 cm) and different weight toothpick containers (approximately 4.50 cm in diameter with a height of 8.60 cm).
### Table 2: Percentage of Children’s Responses to Questions about Objects

<table>
<thead>
<tr>
<th>Group A</th>
<th>Objects are the Same Size</th>
<th>Objects have the Same or Different Weight</th>
<th>Which Object is Heavier?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Percentage</td>
<td>N</td>
</tr>
<tr>
<td>Identical Same Weight</td>
<td>104</td>
<td>86.7%</td>
<td>119</td>
</tr>
<tr>
<td>Non-identical Different Weight</td>
<td>100</td>
<td>83.3%</td>
<td>116</td>
</tr>
<tr>
<td>Non-identical Same Weight</td>
<td>113</td>
<td>94.7%</td>
<td>113</td>
</tr>
<tr>
<td>Identical Different Weight</td>
<td>100</td>
<td>83.3%</td>
<td>120</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group B</th>
<th>Objects are the Same Size</th>
<th>Objects have the Same or Different Weight</th>
<th>Which Object is Heavier?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Percentage</td>
<td>N</td>
</tr>
<tr>
<td>Identical Same Weight</td>
<td>105</td>
<td>87.5%</td>
<td>114</td>
</tr>
<tr>
<td>Non-identical Different Weight</td>
<td>106</td>
<td>88.3%</td>
<td>119</td>
</tr>
<tr>
<td>Non-identical Same Weight</td>
<td>110</td>
<td>91.7%</td>
<td>112</td>
</tr>
<tr>
<td>Identical Different Weight</td>
<td>110</td>
<td>91.7%</td>
<td>119</td>
</tr>
</tbody>
</table>

Average | 106 | 88.3% | 116.5 | 97.1% | 94.8 | 98.8% |

*Notes. N = 120. If the objects had the same weight, participants were not asked which object was heavier.*
3 Table 3: *Comprehension Questions for the Experimental Books*

<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>What was in the bucket that made it heavier?</td>
<td>Toys</td>
</tr>
<tr>
<td>What was dropped from the top of the playground?</td>
<td>Rocks</td>
</tr>
<tr>
<td>What was put into the shoe to make it heavier?</td>
<td>Sand</td>
</tr>
<tr>
<td>What was dropped from jungle gym?</td>
<td>Bucket</td>
</tr>
<tr>
<td>What was dropped from the top of the seesaw?</td>
<td>Shoes</td>
</tr>
</tbody>
</table>
Table 4: Mean Pretest and Posttest Score by Book Condition and Age

<table>
<thead>
<tr>
<th>Age</th>
<th>Realistic Fiction Book</th>
<th></th>
<th>Informational Book</th>
<th></th>
<th>Control Book</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td></td>
<td>N</td>
<td></td>
<td>N</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pretest</td>
<td></td>
<td>Pretest</td>
<td></td>
<td>Pretest</td>
<td></td>
</tr>
<tr>
<td></td>
<td>M (SD)</td>
<td></td>
<td>M (SD)</td>
<td></td>
<td>M (SD)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Posttest</td>
<td></td>
<td>Posttest</td>
<td></td>
<td>Posttest</td>
<td></td>
</tr>
<tr>
<td></td>
<td>M (SD)</td>
<td></td>
<td>M (SD)</td>
<td></td>
<td>M (SD)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td></td>
<td>19</td>
<td></td>
<td>20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.10 (.72)</td>
<td></td>
<td>1.58 (.61) †</td>
<td></td>
<td>1.15 (.99) †</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.80 (1.11) †</td>
<td></td>
<td>2.42 (1.07)</td>
<td></td>
<td>1.40 (1.00) †</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>19</td>
<td></td>
<td>21</td>
<td></td>
<td>20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.65 (.67) †</td>
<td></td>
<td>1.67 (.66) †</td>
<td></td>
<td>1.25 (.72) †</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.10 (1.25) †</td>
<td></td>
<td>2.90 (1.26) †</td>
<td></td>
<td>1.90 (.97)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>40</td>
<td></td>
<td>40</td>
<td></td>
<td>40</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.88 (.72)</td>
<td></td>
<td>1.62 (.63) †</td>
<td></td>
<td>1.20 (.85)* †</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.95 (1.18) †</td>
<td></td>
<td>2.68 (1.19) †</td>
<td></td>
<td>1.65 (1.00)** †</td>
<td></td>
</tr>
</tbody>
</table>

Notes. *p < .05, ** p < .001, †Different than chance (score of 2.00)
Table 5: Participants’ Response to the Open-Ended Questions for the Experimental Books

<table>
<thead>
<tr>
<th>Ranking of Response</th>
<th>Realistic Fiction</th>
<th>Informational Text</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Other</td>
<td>7</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>Book Related</td>
<td>25</td>
<td>24</td>
<td>49</td>
</tr>
<tr>
<td>Objects Fall Down Together</td>
<td>6</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td>Conceptual Understanding</td>
<td>2</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Total</td>
<td>40</td>
<td>40</td>
<td>80</td>
</tr>
</tbody>
</table>
Table 6: *Mean Book Enjoyment Composite Score by Book Condition and Age*

<table>
<thead>
<tr>
<th>Age</th>
<th>Realistic Fiction Book</th>
<th></th>
<th>Informational Book</th>
<th></th>
<th>Control Book</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>M (SD)</td>
<td>N</td>
<td>M (SD)</td>
<td>N</td>
<td>M (SD)</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>3.73 (1.40)</td>
<td>19</td>
<td>3.61 (1.46)</td>
<td>20</td>
<td>3.73 (1.33)</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>3.80 (1.09)</td>
<td>21</td>
<td>3.79 (1.40)</td>
<td>20</td>
<td>2.95 (1.16)</td>
</tr>
<tr>
<td>Total</td>
<td>40</td>
<td>3.76 (1.24)</td>
<td>40</td>
<td>3.70 (1.41)</td>
<td>40</td>
<td>3.34 (1.29)</td>
</tr>
</tbody>
</table>

*Notes.* Book enjoyment is a composite score, both scales ranged from 1 to 5.
Table 8: *Mean Book Comprehension Score by Book Condition and Age*

<table>
<thead>
<tr>
<th>Age</th>
<th>Realistic Fiction Book</th>
<th>Informational Book</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>N</em></td>
<td><em>M (SD)</em></td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>3.70 (0.66)</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>4.35 (0.67)</td>
</tr>
<tr>
<td>Total</td>
<td>40</td>
<td>4.03 (0.73)</td>
</tr>
</tbody>
</table>

*Notes.* Book comprehension ranges from 3 to 5.
8 Table 8: Mean Reading Fluency Score by Book Condition and Age

<table>
<thead>
<tr>
<th>Age</th>
<th>Realistic Fiction Book</th>
<th>Informational Book</th>
<th>Control Book</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N$</td>
<td>$M (SD)$</td>
<td>$N$</td>
</tr>
<tr>
<td>4</td>
<td>18</td>
<td>0.94 (0.24)</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>1.00 (0.00)</td>
<td>18</td>
</tr>
<tr>
<td>Total</td>
<td>38</td>
<td>0.97 (0.16)</td>
<td>34</td>
</tr>
</tbody>
</table>

Notes. Reading fluency score is dichotomous, where 0 = no and 1 = yes.
### Table 9: Mean Age-adjusted Receptive Language (TPVT) Score by Book Condition and Age

<table>
<thead>
<tr>
<th>Age</th>
<th>Realistic Fiction Book</th>
<th>Informational Book</th>
<th>Control Book</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>M (SD)</td>
<td>N</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>106.60 (8.93)</td>
<td>19</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>112.02 (11.62)</td>
<td>19</td>
</tr>
<tr>
<td>Total</td>
<td>40</td>
<td>109.31 (10.59)</td>
<td>38</td>
</tr>
</tbody>
</table>
Figures

1  *Figure 1.* Same weight objects and different weight objects.

Each set has two identical and non-identical objects.
Appendix

1 Realistic Fiction and Informational Story Scripts

Realistic Fiction Book (Words 419)  Informational Book (Words 407)

One day Alice and her friend Luke were at the park. Luke was waving his ice cream around and it fell to the ground! “What goes up must come down!” laughed Alice. “If you let go of your ice cream, it will fall to the ground!” They grabbed their buckets and headed towards the jungle gym.

Luke and Alice climbed to the top of the jungle gym. “I wonder what happens if you drop two things at the exact same time?” Alice asked. Luke said “Our buckets are the same size. But my bucket is full of toys and your bucket’s empty. I think my bucket’s going to reach the ground first. “No way!” Alice exclaimed.

Both Alice and Luke let their buckets go. “Hey, both sand buckets reached the ground at the same time” said Alice. “Let’s go play on the playground” Luke said.

On the playground, Luke took two rocks out of his pocket. “Look!” he said to Alice “Both rocks are the same size, but one rock weighs more than the other rock”. “There are many different types of rocks in the world” Alice replied. Luke dropped the rocks at the same time. “Wow! Did you see that?” asked Luke. “Both rocks reach the ground at the same time, even

What goes up must come down. When an object is tossed up into the air it will always come back down. If you hold out an object and then let it go, it will fall to the ground. That happens even if it’s something you really wanted to keep, like ice cream on a cone!

What happens if you drop two things at the exact same time? Let’s find out how two different objects fall when they are dropped together. First, look at these buckets. These two buckets are the same size. But one bucket is full of toys and the other bucket is empty.

Let’s see what will happen if these two buckets are dropped from the top of a jungle gym at the same time. Down the buckets go! Both buckets reached the ground at the same time. In this example, let’s see how smaller objects fall. Look at these two rocks. There are many different types of rocks in the world. Rocks come in all different shapes and different sizes. Both of the rocks are the same size, but one rock weighs more than the other rock.

The rocks are dropped from the top of the playground at the same time. Look, the rocks are falling down! Both rocks reach the ground
though one is heavier than the other. Just like the buckets” replied Alice.

“I thought heavier things reach the ground before light ones!” said Luke. “Let’s find out” exclaimed Alice. “Give me your shoe please” she said as she took off her shoe. Alice filled up her shoe with sand and left Luke’s shoe empty. “Look, the blue shoe is heavier than the red one”.

“We can drop them from the seesaw!” Luke exclaimed. When Alice was at the top, she held her arms out and let the shoes go at the exact same time! Just like the buckets and rocks, both shoes reached the ground at the same time even though one shoe was heavier than the other.

“What did we learn today?” asked Luke.

“Objects that have different weights and are the same size will fall at the same speed” replied Alice.

Later, Alice explained to Luke “Gravity is the force that makes objects fall to the ground. Gravity affects things that are similar in size in the same way. When objects that are almost the same size are dropped together, they reach the ground at the same time, no matter what they weigh.” Suddenly, Alice’s ice cream fell! “Everything falls down” Luke laughed.

at the same time, even though one is heavier than the other. Just like the buckets.

Some people think that heavier things reach the ground before light ones! Let’s find out if the same thing happens with bigger items, like shoes. In this example, one shoe is filled with sand and the other shoe is empty. Now, the blue shoe is heavier than the red one.

The shoes fall from the top of the seesaw at the exact same time. Look, when the shoes are dropped they fall down to the ground! And just like the buckets and rocks, both shoes reach the ground at the same time, even though one shoe is heavier than the other.

Conclusion. What did we learn today? Objects that have different weights and are the same size will fall at the same speed.

Everything falls down, even ice cream! Gravity is the force that makes objects fall to the ground. Gravity affects things that are similar in size in the same way. When objects that are almost the same size are dropped together, they reach the ground at the same time, no matter what they weigh.