Discovery stories in the science classroom

by

Diana Jaleh Arya

A dissertation submitted in partial satisfaction of the requirements for the

degree of

Doctor of Philosophy

in

Education

in the

Graduate School of Education

of the

University of California, Berkeley

Committee in charge:

Professor P. David Pearson

Professor Dan I. Slobin

Professor Sophia Rabe-Hesketh

Spring 2010

Discovery stories in the science classroom

by Diana Jaleh Arya

Abstract

Discovery stories in the science classroom

by

Diana Jaleh Arya

Doctor of Philosophy in Education

University of California, Berkeley

Professor P. David Pearson, Chair

School science has been criticized for its lack of emphasis on the tentative, dynamic nature of science as a process of learning more about our world. This criticism is the guiding force for this present body of work, which focuses on the question: what are the educational benefits for middle school students of reading texts that highlight the process of science in the form of a discovery narrative? This dissertation traces my journey through a review of theoretical perspectives of narrative, an analysis of first-hand accounts of scientific discovery, the complex process of developing age-appropriate, cohesive and engaging science texts for middle school students, and a comparison study (N=209) that seeks to determine the unique benefits of the scientific discovery narrative for the interest in and retained understanding of conceptual information presented in middle school science texts.

A total of 209 middle school participants in nine different classrooms from two different schools participated in the experimental study. Each subject read two science texts that differed in topic (the qualities of and uses for radioactive elements and the use of telescopic technology to see planets in space) and genre (the discovery narrative and the "conceptually known exposition" comparison text). The differences between the SDN and CKE versions for each topic were equivalent in all possible ways (initial introduction, overall conceptual accuracy, elements of human interest, coherence and readability level), save for the unique components of the discovery narrative (i.e., love for their work, acknowledgement of the known, identification of the unknown and the explorative or experimental process to discovery). Participants generally chose the discovery narrative version as the more interesting of the two texts.

Additional findings from the experimental study suggest that science texts in the form of SDNs elicit greater long-term retention of key conceptual information, especially when the readers have little prior knowledge of a given topic. Further, ethnic minority groups of lower socio-economic level (i.e., Latin and African-American origins) demonstrated an even greater benefit from the SDN texts, suggesting that a scientist's story of discovery can help to close the gap in academic performance in science.

For my students, who first inspired me to venture down Alice's rabbit hole . . .
and for Andy, who followed me the entire way.

Table of Contents

Acknowledgements

I could not have accomplished a dissertation of this magnitude without the help of many people. I would first like to thank my advisor, P. David Pearson, for his constant support and guidance throughout this project. I would also like to thank my other committee members, Dan I. Slobin and Sophia Rabe-Hesketh, who have provided excellent support within their respective areas of expertise.

Many individuals contributed to the successful development of texts and measures used in this dissertation. To my scientists, Adam Mazur and David Doggett, and to my literary expert, Gail Piestrup—you all were invaluable contributors. I would also like to thank the following individuals for contributing to the quality of my study: Lauren Barth-Cohen, Beat Schwendimann, Melvin Piestrup, Susie Goodin, Jennifer French, Susan Ebbers, Marnie Nair, Niccola McClung, Patty Chacon and Anna Nuno.

And finally, I want to thank my family. To my parents, Carolyn and Mahmoud—you provided comfort in many ways, thank you for always believing in me. And to my Aunt Pat—you always seemed to know when to send me hand-knitted "hugs," thank you. And to my father in-law, Terry—thank you for all of your wonderful moral support. And to Andy, the love of my life—thank you for indulging me by listening to all of my stories . . . have no fear, more will surely come.

Introduction

Over the last twenty years, researchers within the field of reading have worked steadily toward building a deeper understanding of student readers and the kind of texts that they should read. A great deal of effort has focused on understanding the fundamental processes involved in reading comprehension (Smagorinski, 2001; Kintsch, 1998; Samuels, 1977; Rosenblatt, 1989), and the particular qualities of the reader that affect these processes (Cunningham & Stanovich, 1998; Velluntino et al., 1996). Reading researchers have also examined text and its role in reading comprehension, which includes issues of readability (Dale & Chall, 1948; Coxhead, 2000), specific elements of coherence and linguistic complexity (Biber, 1992; Nagy & Scott, 2000; McNamara et al., 1996) and issues regarding organizational structures or genre (Donovan & Smolkin, 2001; Duke, 2000; Jetton 1994). The present study stems from this final area of text-based exploration by applying theoretically grounded principles of a particular genre of text to reading comprehension research. More specifically, my interest is focused on the current problems with texts traditionally used in science curricula and on constructing an alternative type of text that both aligns with the cognitive needs of the reader and reflects the actual practices of scientists, the latter of which have been sorely neglected in science education.

Although textbooks have been an integral part of school science curricula over the last eighty years, issues regarding the quality of these texts have slowly gained attention over the past few decades, and more recently science textbooks have been noted to be inadequate in terms of depth, coherence, relevance and accuracy (AAAS, 2002; Niaz & Rodriguez, 2000). Further, science textbooks have taken some of the blame for the lack of knowledge students and teachers have about science as a way of knowing about our world, due to the texts' lack of focus on the tentative, process-oriented nature of science (Bowker, 2005; Niaz & Rodriguez, 2000; Siegel, 1978). It is this problem regarding science textbooks that has inspired me to undertake this study; more specifically I decided to develop a set of experimental texts that would allow me to determine the unique educational benefits of a process or discovery-oriented science text compared to a more traditional knowledge-based text. More specifically, I created two texts that highlight the story of scientific discovery and two comparative texts that convey the same conceptual knowledge but without the story. I developed these texts in order to find out if there is unique value, either in terms of understanding and retention of concepts or interest in those concepts, for the story of scientific discovery.

When we view the process of discovery through a narrative lens, the evaluative stance of the scientist within the context of the scientific field becomes more apparent. Simply, scientists tell stories about their work, which are further evaluated by peers within the respective field of study, and these stories, whether they are published studies or first-hand accounts of discovery, deserve greater attention in science education. Although this idea of narrative as an inherent and important aspect of science is widely supported by scientists and scholars across disciplines (Landau, 1991; Linde, 2001; Mishler, 1995; Ochs & Taylor, 1992; Orr, 1990), there is some contention regarding the notion that stories should be a part of a science curriculum.

This contention largely stems from a view of narrative as fictional tales of fantastic origin, which anthropomorphize animals and simplify natural processes in order to entertain readers at the expense of conceptual accuracy (Duke, 2000, Jetton, 1994; Madrazo, 1997; Mayer, 1995; Rice, 2002). In the first chapter of this dissertation, I present a broader definition of narrative, which is an account of an experience with an evaluative perspective. An "evaluative perspective" in this sense relates to the opinions on the part of the teller. Scientists often tell stories about their work, and these stories include an evaluative perspective based on evidence, which is essentially an informed opinion. In presenting this broader view of narrative, I hope to foster a different conversation about its use in science education, one that focuses on a sense of story that is inherent to science and the use of these stories to make explicit the evaluative perspective of science on the natural world.

The establishment of narrative as a broadly defined construct that encompasses many different types of stories serves as the foundation for the second chapter of this dissertation, which focuses on a particular type of narrative that I call the scientific discovery narrative (SDN). In this chapter, I present my analysis of 12 first-hand accounts of discovery, which revealed four common components of the SDN—love for the work, acknowledgement of the known, identification of the unknown, and the explorative or experimental process toward discovery. I also discuss how these components are absent from traditional educational approaches of the scientific method, which ignore the dynamic, tentative and creative nature of scientific discovery.

The common components of the SDN provided the foundation from which I developed two SDNs and two comparable non-SDN texts. In the third chapter of this dissertation, I describe how I developed these experimental texts based on theoretical and empirical findings about text cohesion, readability and engagement. I also elicited guidance from scientists, literary experts, teachers and students in order to ensure conceptual accuracy, general accessibility and accountability for consistent quality across versions of texts. One of the SDNs presents the story of how Nobel winning chemist Marie Curie (1904) discovered two radioactive elements (titled, *Radioactivity*) while the other tells the story of Gallileo Galilei's (1610) creation and use of telescopic technology to see planets in space (*Seeing at a Distance*). I also created two comparison texts that cover the same conceptual information, save for the story of the scientists. These comparison texts treat the concepts as if they were always known, ignoring the discovery process the led to such knowledge, which is why I name this type of text the conceptually known exposition (CKE). This third chapter is an account of the complex, iterative process of researching, drafting and editing by multiple individuals in developing the SDN and CKE texts to be used in the final experimental study.

In the fourth chapter, I present the experimental study, in which a total of 209 7th and 8th grade students were randomly assigned to read two of the four experimental texts, one of each topic (*Radioactivity* and *Seeing at a Distance*) and each a different genre (SDN and CKE). These participants responded to a series of concept-relevant items immediately after reading each text and an additional series of similar items a week later in order to determine retention of conceptual knowledge over time. In addition to these conceptual recall items, participants also responded to a few interest-related tasks in order to determine any effects of the SDN on interest in science-related topics. The creation and evaluation of measures used are discussed in this paper, followed by results from the experimental study.

Following the fourth chapter, I offer some final remarks that summarize the major learning points presented in this dissertation.

Chapter 1: Scientists as Storytellers: Deepening the Discussion of Narrative for Science Education

Chapter Summary

This chapter introduces the current problem of school science textbooks and the lack of knowledge students and teachers have about science as a process of learning rather than a collection of facts about the world. I argue how narrative, defined as a construct used to convey an experience, is a key element in addressing this problem.

An Introduction to the State of Textbooks in Science Education

Converging work by many cognitive scientists has revealed a lack of understanding about the developmental process of scientific knowledge acquisition (i.e., the nature of science as a way of knowing) among elementary and secondary students and teachers alike (Ackerson & Abd-El-Khalick, 2005; Gallagher, 1991; Griffiths & Barman, 1995; Lederman, 1992; Moss, 2001). In other words, students and teachers possess a fundamental misconception about the development of scientific knowledge, generally viewing it as stable and immutable rather than dynamic, contingent, and situated within the cultural practices of both the larger scientific community and any particular community of scientists working on a common project.

In *The Structure of Scientific Revolutions*, philosopher Kuhn (1962) explains that science is not a collection of absolute knowledge that tells us something true about the world; it is a series of phases in which many scientists contribute new knowledge over a long period of time and these incremental contributions lead to paradigmatic shifts in scientific understanding. Thus, what we think to be true about the world may change depending on the corroboration, testability and logical probability of the amended theory. Teachers and students alike have very little understanding about these features of scientific knowledge and even less about what scientists actually do to gain the knowledge they do have. This lack of understanding can not only foster misconceptions about the nature of science (Lederman, 1992; Ackerson & Abdi-El-kalick, 2005) but can also create a distance between the subject of science and the students who are trying to learn about it (Latour, 1987; Lemke 1990).

In a parallel but related world, studies on science educational texts have revealed that they largely fail to acknowledge scientific processes underlying conceptual information, which may account for the fact that students leave school with little more than decontextualized stores of knowledge accumulated through fact memorization rather than deep understanding of how scientific knowledge develops over time (Bowker, 2005; Leite, 2002; Niaz & Rodriguez, 2000). Thus, my research question:

> What are the cognitive benefits of creating a text for middle school students that reflects the conceptual progression of scientists within a particular field?

I began by analyzing a wide corpus of texts by scientists who explicate their process towards conceptual discovery, and then used this analysis to create two text versions of the same conceptual information, one of which highlights the story of the discovery process and one that does not. Both versions of text were written for middle school students in the United States. Students' understanding of these texts was measured to determine whether there is an added benefit of using the story of the discovery, which I call the science discovery narrative (SDN).

This study makes a unique contribution to the field of reading in three ways. First, unlike many studies that investigate the effects of text genre on comprehension (e.g., Duke, 2000; Jetton, 1994; Rice, 2002), I provide a theoretically grounded definition of the genres used in this study. In particular, I incorporate narrative theory (Bruner, 1990; Labov & Waletzky, 1967; Polanyi, 1985; White, 1980) in the exploration of a specific type of story, one told by scientists when describing their accounts of discovery. This type of story or narrative (i.e., explanation of an event or series of events that expresses an evaluative stance) is one of the many kinds that people create.

Most of the rhetoric regarding the use of stories in science education focuses on a narrow view of narrative and is driven by the attempt to determine whether young readers may develop misconceptions from fictionalized, often anthropomorphized, accounts of important science concepts (Jetton, 1994; Madrazo, 1997; Mayer, 1995; Rice, 2002),. Researchers from both sides of this debate seem to define narrative as fictional tales involving fantastical characters and plot twists and do not acknowledge texts explicating science exploration as narratives. Further, some noted that "acceptable" examples of nonfiction texts contain narrative structure that is often overlooked in discussions about genre and science texts (Ebbers, 2002; Duke, 2000; Jetton, 1994; Pappas, 2006).

Broadening the view of narrative leads to the second unique contribution of this study, which addresses the predominant focus on non-narrative, informational text within middle school science curricula. Explicit discussion of narrative is generally reserved for literature courses that focus on stories with fully developed protagonists and plot twists, but the narrative construction seeps into many facets of human life and work. Thus, I want to provide the opportunity to investigate the narrative structure across disciplinary boundaries, particularly the literature-science boundary (Sohmer & Michaels, 2004), to see what affordances it may provide in characterizing the work of scientists. Finally, in light of the push for teaching the nature of science as a way of knowing as explained earlier, it is important to consider ways in which texts, especially narrative texts, can play a role in this endeavor.

A Definition of Narrative

The following sections present a definition of narrative that is founded on multiple theoretical perspectives and clarified by illustrative texts. My aim is to move from a simple, "single genre" view to a more complex, socio-cultural construct called "narrative" that encompasses many different genres, including those within science. In the most minimal sense, a narrative is a telling of experience that communicates some form of evaluative stance about the teller, the listener and the world.

Moving beyond "literate snails"

When I use the words, *narrative* or *story*, I refer to the human tradition of telling and organizing one's experience (Bruner, 1990; Morson, 1994). This includes the sort of stories one finds in English literature and journalism as well as the telling of everyday experiences (Labov & Waletsky, 1967; Ochs, 1994). It also includes the accounts of scientists in their work and the use of the narrative as a tool for organizing and explicating scientific knowledge with a particular stance toward this knowledge (Bowker, 2005; Sohmer & Michaels, 2004), which is the particular focus of my project. Thus, narrative is not a genre in itself; it is a culturally grounded, psychological construct that surfaces in many different contexts, including science.

The discussion of narratives in science should not be limited to the literary world full of fictitious beings like the talking snail that hitches a ride on a whale's back in *The Snail and the Whale* (Donaldson & Scheffler, 2003). In this story, the whale is suddenly beached due to losing its way, and so the snail saves the day when it uses its secreted fluid to write a message for a classroom of children who in turn call the fire department. The moral of this children's story is simple—even the smallest of us can do great things to save our world. However, many misconceptions about snails and whales accompany this moral. As an easy example, snails are not literate; the shiny trail they leave behind is not used to spell out words. Misconceptions like this serve to reinforce the moral codes of a particular culture (Martin, 1991).

The misconceptions that are introduced for the sake of creating a heart-warming, moral tale inadvertently contribute to further confusion about particular science concepts, but one could hardly blame the authors, whose purpose was to entertain and inspire, not to teach science. However, one must wonder about the consequences for children who are only exposed to texts like this. Many researchers contend that young students are overexposed to narrative texts like the whale story during the early elementary years, leading them to struggle through academic textbooks that are obligatory in later years due to significant differences in organizational and structural features (Duke, 2000; Donovan & Smolkin, 2001). A possible conclusion, then, would seem to be that narrative has a debilitating effect on science learning, as it is not involved in the work that scientists do nor is it present in scientific texts.

However, narrative is more than the fantastical, whimsical literature that pervades the early elementary curricula. In order to fully understand the meaning of "narrative text," it is critical to move beyond the surface features of anthropomorphized characters and interesting plot twists. It is equally important to have a broader notion of "text" that includes both written and oral discourse. Thus, I address narrative as a universal construct that will encompasses the many different types of stories, whether spoken or written, that people create for a variety of purposes in a variety of contexts.

Theories of narrative

There are many different kinds of narratives, used for many different purposes. The following discussion presents the broad landscape of theoretical perspectives on narrative along with exemplary texts.

Approaches to studying narrative. There are two general approaches to the study of narrative. One approach involves a "structural" account of stories. According to this view, narratives consist of components that indicate a story type. Literary scholars (Forster, 1927; Leitch, 1986; Prince, 1973; Propp, 1928) use this approach to examine the common elements of folklore and novels. Psychologists, on the other hand, use this structural approach to identify the components of an "ideal" story that support accurate memory recall (Mandler & Johnson, 1977; Rumelhart, 1977a; Thorndyke, 1977). Sociolinguists (Labov and Waletzky, 1997; 1967) have a different focus from the previous two groups in that they are more interested in how people narrativize their experience (i.e., create narratives), thus using a structural lens to analyze the common components that make up this everyday practice of storytelling and the ways in which this ability develops over time (Fivush, 1994; Hudson, 1990; McCabe & Peterson, 1991).

A second approach to studying narrative focuses more on the culturally embedded functions of storytelling, and the ways that humans make sense of their experiences and relate those experiences to others, thus shaping their cultural identities. Psychologists and sociolinguists who use this approach (Berman, 1995; Black & Wilensky, 1979; Bruner, 1990; Polanyi, 1985) think that the unique, human qualities demonstrated by narratives cannot be solely defined by the actual form that they take but should be defined more on the effects that they have in creating and reaffirming one's cultural identity. These two general approaches are not mutually exclusive in that both contribute to a foundational framework for a deeper discussion about the role of narrative text in science education. Further explication of the theoretical perspectives introduced above will provide a lens through which different types of scientific texts can be examined and reconsidered as different types of 'narrative' science texts.

The boundaries of narrativity. A beginning point for unpacking the definition of narrative is to explore what falls outside the boundaries of a narrative. The following description of photosynthesis illustrates this point:

> A plant increases in the number and size of leaves and stems as it receives energy from the sun through photosynthesis. Photosynthesis is the process where the green pigment in the plant's leaf (chlorophyll) absorbs energy from sunlight and, using this energy, water, and carbon dioxide, produces oxygen and simple sugars. The plant then uses these sugars to make more complex sugars and starches for storage as energy reserves, to make cellulose and hemicellulose for cell walls or with nitrogen, to make proteins. How the plant uses its energy depends on the developmental stage of the plant and on environmental conditions ("Plant growth and development as the basis of forage management," 1993, para. 2).

This text falls short of being a narrative in that although it describes an event (in this case the process of photosynthesis), there is no indication human perspective or evaluation about this event. The process of photosynthesis is merely reiterated as a series of actions that occurs every time a plant produces a new leaf and not reflected upon. According to English professor and literary scholar Leitch (1986), what makes story a story goes beyond the causal and temporal sequences of events, which is why structural aspects alone cannot fully explicate the meaning of narrative. Leitch emphasizes the crucial element of "tellability," defined in the same spirit as psychologist Bruner (1990) who views stories as having engaging and "conventionally uncanny" qualities that make the story worth reading about or listening to. Plus, there needs to be a point to the story; a story worth telling is a story that has something to say about the state of affairs.

Linguist Polanyi (1985) argues that even the most informal, casual story has a distinctive structure that makes a point about something in the world. This point is generally some form of opinion or critical evaluation that has some form of dramatic impact on the receiver. The drama stems from an unusual occurrence that is meaningful in some way. This is what separates stories from mere reports that recount an event with no particular significance. The passage above has no clear tellibility in that no particular point is being conveyed other than to merely define photosynthesis. One could imagine this passage to be read by a student, who in turn may need to explicate (some may even use the term, "regurgitate") the process of photosynthesis for a homework assignment or test. This is not to state that such texts are unnecessary; one could also imagine an individual who is struck by the wonder of plant growth in her garden and would like to know more about the actual process, making this description just as relevant as any story her friends may tell her within the course of the day. The point here is that although a description about a natural process like photosynthesis can be useful and relevant to the needs of the reader, it is not a narrative.

Narrative is more than a telling of an event, but it is also more than just presenting an evaluative stance. Consider the following text (Spak, 2009) about a new way to foster plant growth:

> For growers and gardeners of all kinds, Smart Grow mats could be revolutionary. Tucked around the stems of plants, the cheap, natural mats keep weeds and pests away as they spur growth. There's just one thing holding them back: They're made from human hair. "No matter what you do," says CEO Blair Blacker, some people just "won't get over it."
> Blacker's warehouse is stacked with roughly $400,000 worth of hair. It doesn't smell great, but that's "the odor of money, as far as we're concerned," he tells Marketplace. The Florida company had sales of $300,000 last year, and it's looking for the capital to go national. If that happens, it could revolutionize the agriculture industry: One farmer says the product saves him $3,200 a year ("Human hair grows plants, profits for Fla. Company" 2009).

This brief report does not describe anything happening across a given time period. Instead, the author presents an argument that using human hair for fertilizer is a lucrative idea, a point that is confirmed by a company chairman and a farmer. So, here is a "point" without the telling of any process or event, without any sense that time has passed.

Narratives as historical records. Polanyi (1985) defines narratives as "kinds of discourse organized around the passage of time in some 'world'" (10). Philosopher Ricoeur (1980) points to the construct of time as a critical feature in narratives; it is in this sense of moment-to-moment framing of a story that what actually happened becomes restructured as a relevant telling in the present moment. The following excerpt (Akinwande, 2009) about the historical development and use of plant fertilizer illustrates this feature of time:

> With time, natural fertilization became more refined. For instance, the ancient Egyptians added ashes from burned weeds to soil. Other materials used in ancient times included sea shells, clay, and vegetable waste. Starting from the early 17th century, people researched other modes of fertilization, particularly those of a chemical nature. For instance, German-Dutch chemist Johann Glauber (c. 1604 to 1670), developed the first mineral fertilizer, which comprised saltpeter, lime, phosphoric acid, nitrogen, and potash ("The history of fertilizers," para. 2 & 3).

The excerpt above certainly signals a passage of time, yet the evaluative stance, if one exists, is considerably subtle compared to the persuasive report about hair fertilizer. Does this excerpt represent a narrative, then? Literary scholar White (1980) distinguishes between two terms, "narrate" and "narrativize," which he presents as two different ways of recapitulating past events. According to White, historians need not tell about the past in narrative form. They can simply report the events in realistic terms with no significant point made about the event itself. This is what White considers "narrating" rather "narrativizing" because it does not convey any meaning other than what took place. When storyteller narrativizes, she creates a story of a particular happening in relation to her view about herself and the world around her.

Considering that no specific "point" is being made within the historical excerpt about fertilizer, it would not seem to be a narrative. However, Ricoeur's (1980) definition of narrative seems to disagree with White's dichotomy. Ricouer characterizes narrative discourse as essentially a restructuring of events as they actually happened. We tell about events that occurred in relation to the present moment and the cultural values of today, which is the reason one can never know historical events as they actually took place. The evaluative stance in this sense is the selection of what parts of history should be told or even emphasized, bringing a significance to those parts of the past that are the most relevant to the present time and culture.

Some critics would call this a rather charitable interpretation of historical texts, and point to the "white washing" of past events for the sake of promoting particular ideologies (Anderson, 1971; Altbach, 1991; Apple, 1991; de Castell etal., 1989; Ivanič, 1998). It is the reconstruction of historical time into human context that creates the narrative, and depending on the extent of this reconstruction, more or less of the "truth" about what took place may be known. However, no historical text can reveal the absolute truth about the past, as literary scholar Morson (1994) explains: "Life as it is lived is not storylike, and so we may suspect that whatever story we choose to tell about it will alter it. Life includes all sorts of extraneous details leading nowhere, but good stories do not. Narratives are more successful if they display a structure, which is hard to find in life" (pp. 19-20). Based on this perspective, all historical texts should be viewed with a critical eye and treated as narratives that have more to tell than past events. For example, the final paragraph connected with the fertilizer excerpt above reveals a modern agenda:

> Although organic fertilizers are still used today throughout the world, chemical fertilizers are more popular. Also, research is still being conducted---to reduce the harmful environmental effects of fertilizer use, as well as discovering new, less costly sources of fertilizers ("History of fertilizers," para. 6)

Narratives as cultural texts. Terms like "organic," "environmental effects," and "less costly" are a part of the current discourse in western society. Presenting these terms anchors the past in a way that creates a timely relevance and importance for the reader. Polanyi (1985) called this anchoring the "cultural text" of narratives (i.e., shared norms within a particular culture). A cultural text is any mode of meaning that would be familiar to a group of individuals that share a particular set of culturally defined beliefs. This is similar to Bruner's (1990) concept of folk psychology, which he describes "as a system by which people organize their experience in, knowledge about, and transactions with the social world" (35). Both cultural text and folk psychology include the knowledge of shared norms about what is worth telling. Thus, an essential quality of narrative is its reflection of the cultural values and ideals of the author who is telling about the past.

The cultural anchors of narrative did not go unnoticed by cognitive psychologists like Mandler and Johnson (1977), who found that the less a story resembles the 'ideal' format with all conventional features, the more likely that information within the story will be forgotten, misinterpreted, or misremembered. These researchers called a person's ideal format "schema," a term first defined by the renowned social psychologist Bartlett (1932), who documented how participants, when recalling the events of a story, omitted or altered particular aspects according to prior expectations or, rather, their schematic framework. In this sense, schema is a system of knowledge one elicits and uses when making sense of the world. For example, if an individual read a story about a spirit, but did not have any cultural understanding about this concept, the memory of this point in the story will be distorted or omitted. Bartlett found that these alterations resulted in stories that were more conventionally congruent with the storyteller's cultural schema, revealing that as time increases, one is able to restructure an account that is more familiar and meaningful.

In light of the obvious power that narratives have in distorting past experience, or even experience as told in a fictional account, one must ask the question, should narratives be avoided, especially when we try to remember historical events? What use is narrative if it distances us from knowing the "truth" of what happened? The point that Ricoeur (1980) and Morson (1994) try to make is that this "distortion" that is essentially the cultural text, folk psychology or schema of a narrative, is unavoidable. In our attempts to make sense of an experience, we change it, and the important lesson is to acknowledge this fact. Only in understanding the ways in which narratives construct our sense of reality can we truly appreciate how telling stories brings people together, as well as pull them apart.

Everyday narratives. One of the most ubiquitous types of narrative is the everyday storytelling that people share with one another for a variety of reasons. We reminisce about the "good ole days" at the local bar, around the water cooler or at the barbershop. We explain what "really" happened to the news reporters, jury or school principals. We even like to (or in some cases, have to) tell each other about how we spent our time apart. I have written the following example to illustrated this point:

> Honey, I'm glad you're back. I had a fairly productive week. I read about this new way to fertilize plants and I tried it out. Get this—they use human hair as fertilizer! It wasn't that expensive either, just a little creepy using somebody's hair. I just packed the hair mats in around the flowerbeds, so we'll see if it makes a difference. If it works, I might continue to use it, but don't worry, it didn't cost that much.

In his book, *Acts of Meaning*, Bruner (1990) explains the central role of narrative in people's construction of "meaning" about their world and about their own identity. Through everyday storytelling, people reaffirm who they are in relation to others and their place in their cultural context. From the brief story above, we are able to understand that the teller likes to garden and considers herself somewhat of a bargain hunter, even if these bargains require less-than-appealing materials. We also get a sense that the garden, however large it is, belongs to her and not the listener, who we assume to be the spouse waiting for the final price of this "fairly productive week." Thus, narrative is an act of shared meaning. The root of this culturally shared meaning is the previously mentioned construct, "folk psychology," which Bruner defines "as a system by which people organize their experience in, knowledge about, and transactions with the social world" (p. 35).

Folk psychology helps us to identify ourselves as well as others within our culture. It is the knowledge about our cultural norms that we use in determining that "things 'are as they should be'" (p. 40). If, for example our lady gardener above told her spouse that she went to buy human hair (not necessarily a wig) with no other explanation, the conversation would be left in an odd position. Why would she possibly need to buy human hair? Narratives help us understand non-canonical behavior via the protagonist's intent or desire. The gardener sets up her story by orienting her listener to the purpose of finding a new cost effective fertilizer. This orientation provides a frame for understanding the non-canonical behavior of buying hair. Narrative is one of the most powerful tools of folk psychology for this reason, creating a connection between the intent of the protagonist and the natural happenings of life. This is the very reason why truth is stranger than fiction; without a story, one cannot make sense of the extraordinary, which, ironically, happens all the time. I have written the following everyday narrative:

> I remember this one small rhododendron shrub in my backyard that kept dying on me, no matter what I tried. At first I thought it was in the wrong place, not getting enough sun. But then one morning, I watched my dog as I let her out in backyard. She ran right to that shrub and squatted down. Mystery solved—kill the dog.

Linguists Labov and Wiletzky (1967; 1997) conducted an extensive analysis of English oral recapitulations of past experience much like the one above in order to determine the basic units that consistently compose what they consider to be a complete, simple narrative. They argue that "fundamental structures [of narrative] are to be found in oral versions of personal experiences: not the products of expert storytellers that have been retold many times, but the original production of a representative sample of the population" (1997, p. 3). After analyzing six hundred tellings, the researchers decided on a clausal definition of narrative that is described as "(a)ny sequence of clauses which contains at least one temporal juncture" (1967, 28). This sequence includes an orientation (a small shrub in the backyard), complicating action (shrub keeps dying), evaluation (confusion about the plant; frustration with the dog), resolution (dog squats on poor plant) and possibly a coda (joke about killing dog).

Of particular importance for Labov and Wiletzky's analysis are the temporal sequence of events that matches the order of what happened and the evaluation, which helps the audience to understand the significance of the narrative. This structural account is similar to the "narrative superstructure" presented by psychologists Kintsch and van Dijk (1983) that was used a way to learn more about cognitive processes. More about their analytic approach will be described in the following chapter, as it served as the central framework for analyzing first-hand accounts of scientific discovery.

Although the narrative sequence presented by Labov and Waletzky (1967) seems to fit the example above, not all tellings necessarily highlight a problem or complicating action. The earlier telling of the wife using hair fertilizer shows no signs of a particular problem in that there is no indication that she's having any particular issues with her garden. Although Labov and Waletzky consider the complicating action as the main body of the narrative clauses in their database, it must be pointed out that in all cases within their analysis, participants were prompted with a series of questions (e.g., 'tell me about the most troublesome girl in your neighborhood') that may have encouraged the telling of narratives involving conflict. In response to the question, "tell me about your day," a complicating action might be present in the responding narrative, but it would not necessarily be expected.

The "passing of time" in narratives. Narratives do not always highlight problems to be solved, and neither do they always reconstruct past events. Most of the narratives we read in the form of fictional novels and fairy tales are evidence for the expansive imagination of authors who write about magical worlds in other dimensions of space and time. Moreover, many of these published stories are written in present tense, giving the effect that the reader is right next to the protagonist as a story unfolds. Thus, narratives do not necessarily describe an experience that had already taken place. What seems consistent is a passing of time, regardless of the relationship between the time of the experience and the time of the telling. In fact, some narrative theorists and researchers have focused on stories that point to future experiences. I have written the following example:

> Remember when we were at the hardware store yesterday? We agreed that while you're away next week, I would focus on the garden in back and buy new plants to fill up that one corner next to the tool shed.

Linguistic anthropologist Ochs (1994) focuses on the "becoming" aspects of narrative in her work that describes how recapitulations of the past also project future experiences as illustrated in the example above. She writes, "One or another teller may see the point of the story to include what certain past events mean with respect to their own or others' future experiences. A sense of the future may be fundamental to the design of the past events from the very beginning of the narrative . . ." (p. 107).

Ochs explains that often it is the future projection of the teller that carries the point or significance to a story about a past event. She examines a series of texts containing "conversational stories" and their indices of future time. These stories are drawn from a previous study involving dinner conversations among twenty white English-speaking families. One particular example describes a father being reminded by one of his children that he promised him chocolate after dinner, referring to a future event. The child's recollection of the father's promise indexes the future event of eating dessert, which is the child's present concern. Ochs argues that this example "supports the notion that future time can play a major role not only in warranting but in *structuring* stories of past experience" (p. 119).

Educational researchers Sperry and Sperry (1996) expand on this notion of future time with their finding of "hypothetical narratives," (i.e., stories told about events yet to happen) in narrative-like conversations that emerged naturally from mother-child dialogue. These researchers focus on how this narrative-like talk can represent imaginary events, not only events in the past. One example illustrates how a toddler tells his mother about a rat that lurks in the bushes outside. The mother indicates that there is no rat in the bushes, but the child insists that a man will get the rat that is in the bushes. Sperry and Sperry explain that the child has created a fictional narrative about a rat and a man that will catch it. The point of this 3-year longitudinal study was to capture the naturally occurring narratives of African-American toddlers, which tend to be more fictional (imaginary events) in present time rather than temporal accounts of real past events. Are such hypothetical narratives limited to young children? Is there no purpose for such storytelling in adults?

> I'll just bet that if you plant that rose bush near the tool shed, aphids will show up from Fred's backyard and set up camp over on our side wreaking havoc. Then you'll have to find something more than human hair to save those roses.

We can imagine hypothetical narratives like the one I constructed above in everyday conversations among adults. Regardless of age, people use narrative to anticipate hypothetical events, thus as a type of dialogic tool for problem solving (Ochs, Smith & Taylor, 1989). This notion of future events in narrative will be revisited within the next section, which addresses narratives in science.

Based on the theories discussed in this chapter, narrative is a culturally bound and biased human construction that describes an experience. Although a narrative indicates a passage of time, the events, whether imaginary or real, do not necessarily take place in the past. The constant elements, therefore, are the temporal sequence of experience and the narrator's evaluative stance on this experience. And being that narrative is a human construct, it seems logical that narrative would play some role in science, which is the human endeavor of knowing about the inner workings of our world and its place in the universe. Thus, the following discussion will address and document the ways in which narrative is and can be used to convey scientific understanding.

Narratives in Science

Are scientists storytellers? If experiences that are not completely forgotten are inevitably reconfigured into culturally meaningful and relevant narratives, then this should apply to scientists as well, who often write about their experiences of creating or discovering new knowledge about the world. Scientific studies could therefore be viewed as a particular type of narrative that describes a replicable process through which new scientific knowledge was attained. Certainly the cultural relevance and evaluative stance is present in such publications that invariably explicate the purpose and necessity for a given study, as illustrated in the following excerpt from a study on plant growth by ecologist Abrami (1972):

> Temperature is the most important short term variable controlling plant development and growth. Fluctuations during the day-night period are of physiological importance as are the seasonal or annual changings in temperature.
>
> It has been difficult to adequately relate seasonal temperature fluctuations to those of growth. Effects of temperature occurring in plants are much more complicated when one considers the organism as a whole than appears from studies of a single biological reaction (Went, 1953). The problem appears even more complicated if one considers the interaction of other environmental factors, notably light (from, *Ecology*, vol. 53, no. 5, p. 893).

In this introductory text, Abrami establishes the importance of learning more about the effects of temperature and light on plant growth and sets the stage for his study on the intersecting variables affecting such growth. The first paragraph in this excerpt seems to have an introductory function similar to what Labov and Waletzky (1967) called the orientation, which generally introduces the major players in the told action, including the time and place. However, instead of introducing people, scientists are introducing concepts, like temperature and its importance in a given problem. The orientation is the first attempt in motivating the listener or reader to care about the subject of the story, and in this sense the scientist must at the beginning provide a good reason for why the reader should continue to read about this study. If the scientist can establish enough "buy-in", then the reader finds value in the rest of the story.

The central problem, or complicating action, begins to surface in the second paragraph, where Abrami signals his evaluative stance on the subject by claiming the field's general lack of knowledge about immediate, interacting effects on plant growth. Once his prerogative is established, Abrami carefully describes all the materials and procedures involved in a systematic study of climatic conditions for seven different species of plants:

> In the present study Lindsey and Newman's method was modified by using two temperature thresholds. I found it necessary, as did Robertson (1968), to consider a possible different lower temperature threshold and an upper threshold for every growth period of 24 hours . . . . In order to choose the most suitable sample site of growth and temperature units, the row pattern of growth and temperature for the periods of study was examined. Fluctuations in temperature are not periodic, however a period of seven days

> was chosen as not too short and not too lengthy to describe the most common type of temperature variations in the climate of Padua (pp. 895-896).

Abrami carefully outlines the conditions and methods used in his experiment, supporting his choices with established methodology of other researchers as well as his own knowledge about the site of the study. The choice to use a particular framework over others could be considered as an evaluative stance, especially if there are competing frameworks available to the scientist, which is often the case. In this particular instance, Abrami explains why it was important to alter what seems to be a well-established framework in the field of ecology. There is no further explanation of the "Lindsey and Newman method" other than what is presented, suggesting that the intended reader knows this method quite well. The fact that Abrami justifies his modification of this method by citing the work of other scientists suggests that readers may find the modification suspicious or confusing otherwise. And once the data are analyzed, out comes the "resolution" of the story:

> Two points are evident from the results of the calculations of correlation. First, the level at which heat energy is required for stem elongation was greater at the end of the process. This trend occurred except for a period referred to previously as the concave portion of the growth curve which corresponds to an important stage in the reproductive process . . . . The second point is the lack of a positive correlation between growth and temperature during some periods of stem elongation. Morphological developments such as leaf and branch initiation which may modify the rate of stem growth cannot explain the phenomenon completely. At the same time, as already stated, we can exclude any effect of light and soil moisture (p. 899).

The initial problem of insufficient information about the variables affecting plant growth is somewhat solved by the conclusions drawn from observable data. Based on a series of choices made during the experimentation, Abrami provides the two take-away points for the reader to consider. So then, what happens next? Often what is viewed as a call for further research can, in a sense, be the "coda" of the study: "Other causes are indicated which impinge on the temperature responses of the plants and which would have to be determined by further investigations" (p. 899). This final statement signals to the reader that the study is not the final word on this issue, but merely a stepping-stone for studies to follow.

Why view studies as stories?

Published scientific studies like the one examined above could be considered a particular type of narrative, but an important question is whether or not it *should* be considered a narrative. What do we gain from seeing these published studies as stories? I believe that the same argument for seeing historical texts as narratives about the past applies here. If we approach a scientific study as though it were a narrative, then we are primed to see something more than what might otherwise be considered a disinterested, objective, systematically generated data that describes something "true" about the world. We have a greater opportunity to see the evaluative stance that reflects the values and opinions of the scientist and the shared cultural norms of a given field of work. The introductory paragraph of a study more than thirty years after Abrami's publication illustrates how current issues of climate change have supplanted earlier, less thematically based investigation of specific environmental phenomena by Clelend et al. (2006):

> Shifting plant phenology over the past several decades provides compelling evidence that natural ecosystems are already responding to human-caused environmental changes (1-8). Earlier flowering (3-5) and earlier peak in primary productivity in satellite data (6, 7) in the northern hemisphere in recent decades are correlated with rising temperatures (8). Experimental warming leads to earlier flowering (9, 10), but data on the phonological impact of other cooccurring global changes are limited. Additionally, previous experimental studies have not linked species-level observations of flowering date with remote sensing-based measures of the phenology of ecosystem primary productivity (from "Diverse responses of phenology to global changes in grassland ecosystem" *PNAS*, vol. 103, no. 37, p. 13740).

The authors orientate the reader to the evidence that people are affecting plant growth, with significantly more citations to support this claim than what Abrami (1972) offered in his beginning point about temperature, which is a more observable variable than human influence on plant growth.

The earliest references presented by Cleland and her colleagues to support their strong statement above are from the late 1990's, thus showing "the past several decades" to actually mean over the past two decades, at most. This entire introductory paragraph has more references than what Abrami cited in his entire article, which suggests that the human influence in changes of plant growth is a debatable issue. It seems as though the subject of plant phenology has become more organized around a central, thematic narrative that reflects a current cultural concern. Taking notice of such narrative devices used, albeit inadvertently, in publications like these help us to understand the underlying purpose and influences of the scientist, who is doing more than just conducting experiments and reporting the findings. Rather, the study fits into a larger story, and it is only in the context of that story that we can fully appreciate the meaningfulness and usefulness of the study.

Perspectives of Narratives in Science

Paleoanthropologist Misia Landau (1991) explains in her book, *Narratives of Human Evolution*, that when a scientist attempts to compose a response to a question that begins with the word "why," she is invariably composing a narrative. Sequences of events in the past are reconstructed for the purpose of creating a cohesive account about the world in which we live. Landau focuses on the concept of human evolution and how scientists like Charles Darwin depended on the narrative framework as much as they did on evidence to present a compelling account of human evolution. Landau refers to Propp's (1928) functional analysis of the heroic epic as a framework for examining scientific writing on human evolution from the nineteenth and twentieth centuries. Landau discovered a similar story structure across these essays; the developed being (man) is able to transform and triumph (against the wilds of nature), thus shaping human civilization as we know it today.

Accounts of human evolution tend to vary according to the degree of emphasis on the development of terrestriality (movement from trees to the ground), bipedalism (walking upright), encephalization (brain development) and civilization (group dynamics and order). However, regardless of which of these four concepts is highlighted, there is a similar story in play; the developed being is able to transform and triumph, thus securing human civilization as we know it today. The transformation occurs once the hero is bestowed a gift from a donor. In evolutionary accounts, hidden agents like natural selection or genetic factors play the role of donors that provide gifts of intelligence or tools, which lead to transformation. This account resembles Propp's functional structure of Russian folklore. A final function of the traditional narrative structure, according to Propp, involves the hero being retested after transformation. Folktales often depict the hero faced with a great challenge that will lead to her ultimate triumph. Similarly, narratives of human evolution mention that what was a human's greatest weapon, civilization, will ultimately become the greatest threat.

Landau is quick to add that this comparison of scientific work with archetypal narrative structure is not a criticism of the field but an admission that scientists inevitably tell stories while learning about the world. Landau argues that narrative is not only a tool for helping readers to understand scientific phenomena, but that students within a scientific field like paleoanthropology must embrace their role in the storytelling tradition as demonstrated by Darwin and others.

Landau's argument for narrative in science is echoed in studies by Linde (2001) and Orr (1990), both of whom have identified narrative telling as an essential tool for recalling and organizing knowledge among applied scientists. Anthropologist Orr (1990) studied a group of photocopier technicians and how they use narrative, again inadvertently, in accomplishing the task of repairing machinery for their customers. She explains that the scope of the technicians' repair work is much broader than just fixing or replacing equipment; their success in addressing the needs of their customers depends heavily on their ability to create and use elaborate narratives that are shared among their group to not only address current problems but to anticipate future problems as well. These narratives are elaborate in that they contain details that may seem unimportant but may prove invaluable in future issues with other customers. Orr writes, "Accordingly, the technicians' stories are full of insignificant details, instructions for practice in the form of short cuts and fixes, and new ways in which the relationship of customer, machine and technician can take paths never anticipated" (170-171).

The shared narratives of the technicians in Orr's study are founded on a social network that involves a relationship between the machines and the customers who use them. Orr explains that each machine is an integral part of a social setting, and that the tasks involved in repairing the machines are best framed as repair and maintenance of the social setting itself. Thus, the narratives shared and used within this field of work provide a context for which the technicians are able to work efficiently to maintain the social network by efficiently addressing the needs of the customers, whatever the final outcome be repair or replacement. Orr describes this use of story-telling both as a "preservation" of knowledge as well as a tool for "subsequent diagnoses" (178).

NASA Senior Research Scientist Linde (2001) agrees with Orr that narratives occur informally within a particular group for the purpose of efficiency and future accomplishment, but she also argues that institutions should also take care in creating opportunities that will allow for such stories to be recorded and used widely so that more can profit from such valuable knowledge. Linde views narrative as a powerful type of "tacit knowledge" that acts as a vehicle through which explicit knowledge can be conveyed to members within a particular group with shared goals (160). Tacit, or implicit knowledge is knowledge that one does not know explicitly but uses implicitly, such as the skill of riding a bike. For those with the skill, it is much easier to ride a bike than to explain how to ride a bike. Likewise, it is much easier to tell a story than to explicitly explain what a story is, as evidenced by the theoretical discussion presented earlier. Similar to Orr's argument, narratives within a given professional group communicate more than just past events. Narratives also provide a way to express group identity so that members not only know important content, but they also know the virtues or beliefs characteristic of being an ideal member of the group.

Linde argues that professional institutions should embrace the storytelling that occurs within fields of particular expertise by creating "opportunities for kinds of interactions which allow certain kinds of stories, or even specific highly valued stories to be told" (165). Such occasions may include review meetings in which individuals within an institution are able to reflect on past experience and share incidences of what went well and what did not. Conversations at the water cooler are also valuable opportunities for appropriate narrative propagation rather than for idle gossip. Linde's point is this: institutions must "provide occasions for storytelling, and even more effort to capture, record and make appropriate stories available in a usable and credible form" (170).

Psychiatrist Mishler (1995) also makes a direct connection to narrative and science texts in his in-depth analyses of numerous science texts across various scientific fields. He concludes that narrative is a "legitimate" and necessary device in scientific research. He argues that the positivist stance that has historically dominated the sciences is shifting in light of the influx of narrative research (much of it presented here). This evolving perspective acknowledges the point that "we 'story the world,' making meaning of different events and experiences through our tellings and retellings of stories for different purposes in various contexts" (117). Linguistic anthropologists Ochs and Jacoby (1997) extend Mishler's argument by showing how narrative is present even as scientists talk with each other. They found in their analysis of interactions among physicists that a significant portion of time is dedicated to reaching consensus on experimental events which ultimately lead to presentations and publications. Thus, these researchers view the creation of a scientific study as a co-constructed narrative that helps us understand a particular phenomenon of our world. This process of co-construction has been compared with the problem solving narratives among parents and children during dinnertime conversations (Ochs & Taylor, 1992; Ochs, Smith & Taylor, 1989). This is not a claim that every instance of reaching consensus is an act of co-narrativization; the focus of composing ideas into a future publication is one way that scientists are able to identify the most salient and relevant points from experimental or exploratory data.

The act of summarizing major findings from data becomes more narrativized when scientists begin to think about how to present these findings to potential readers. As a discussion turns to publication, certain points from the data may surface as more universally interesting and salient. The central issue of global warming for Clelend et al.'s (2006) study described earlier, for example, guides the contributing authors as they discuss their data. Thus, the co-construction of scientific publications is a way to build shared meaning about a given process of scientific investigation.

Conclusions

At the beginning of this chapter, I pointed out two outstanding problems about science education. First, students and teachers in science classrooms have insufficient understanding about science as a process and an evolving body of knowledge. Second, science textbooks provide little support in rectifying this first problem. I linked these issues with the current debate about narrative in science texts by arguing for the consideration of a broader, more complex view of stories and the exploration of the ways in which narrative is, and should be, an integral part of texts produced in science. Narrative is a human construct that people use, intentionally or not, to tell about many different types of experiences, some of which are discussed in this chapter. In the following chapter, I further explore one particular type of narrative in science, which I call the "science discovery narrative." I review first-hand accounts of discovery and discuss their potential benefits for improving the quality of school science texts.

Chapter 2: Bringing "real science" into school science texts: An exploration of first-hand accounts of scientific discovery

Chapter Summary

This chapter is an account of the first steps in the development of a prototypical school science text that I call the "science discovery narrative" (SDN). The SDN serves as one of two versions of text used in a study comparing interest levels, short and long-term scientific understanding and recall of U.S. middle school students for two different genres of science texts, which I describe in the fourth chapter. This second chapter will describe the methods for analyzing 12 first-hand accounts of scientific discovery, which revealed four common elements of the SDN—a love for the work, the known, the unknown and the exploration process, all of which demonstrated clear differences from the more common scientific publications.

Introduction

School science textbooks in the United States have been widely criticized for their lack of meaningful content as well as for their lack of attention to the historical evolution of scientific processes underlying conceptual information. Two extensive studies conducted by the American Association for the Advancement of Science (2002) revealed that many scientists and science educators viewed the top selling elementary science textbooks as sub-standard in general quality of information and in the communication of the major ideas in science. An even more extensive text analysis by Niaz and Rodriguez (2000) has shown that secondary textbooks created since the late 1920's have ignored the historical progression of science as a body of tentative, refutable knowledge, leaving the readers to think of science in terms of facts and principles rather than the pursuit of a better answer for a given problem or question about our world. The absence of such discussions in most contemporary science textbooks may partially account for the fact that students leave school with little more than decontextualized databases of knowledge accumulated through fact-memorization rather than deep conceptual understanding (Bowker, 2005; Niaz & Rodriguez, 2000; Siegel, 1978). Many cognitive scientists have echoed this argument by pointing out a lack of understanding about the evolutionary process of scientific knowledge (i.e., nature of science as a way of knowing) among elementary and secondary students and teachers alike (Ackerson & Abd-El-Khalick, 2005; Gallagher, 1991; Griffiths & Barman, 1995; Lederman, 1992; Moss, 2001).

Although the current sub-standard conditions of American school science textbooks served as the impetus for the present analysis, the problem with school science textbooks is not limited to the United States. Educational researchers around the world have attempted one of two general approaches for improving the quality of learning through school science texts. While one approach focuses on instructional strategies for facilitating comprehension of current science textbooks (e.g., Glynn & Muth, 1994; Unsworth, 1997; Vacca & Vacca, 1998), the other approach explores alternative sources for science texts (Baram-Tsbari &Yarden, 2005; Taşdelen & Köseoğlu, 2008; Wellington & Osbourne, 2001). There have even been attempts to create science literacy programs that address both instruction and alternative texts (Guthrie et al., 2004; Palincsar & Magnusson, 2001). I believe that both approaches (or any combinations thereof) are necessary in addressing the overwhelming need for improved methods and texts for school science curricula. My contribution to this effort resides within the second approach of exploring alternative science texts. In particular, I call into question the absence of the evolutionary, tentative nature of science in school science textbooks and take to task the seeming lack of purpose and meaning within these texts. More specifically, I investigate first-hand, written works of scientific exploration as a foundation for creating texts that highlight the true nature of science as a way of knowing about our world.

First-Hand Accounts of Scientific Discovery

Investigations into the advantages of supplanting traditional science textbooks with adapted versions of primary literature by scientists (i.e., scientific studies and research articles) have gained attention in recent years. Baram-Tsabari and Yarden (2005) found in their study of high school students with varying levels of prior knowledge that those who read information based directly on the scientist's study demonstrated stronger levels of inquiry methods (i.e., science as process) compared to those who read the standard school texts. These authors attribute the difference in inquiry-based skills to the principle that "the scientist who did the research is also the one describing it in the article," which "closes the gap between public knowledge and the frontiers of scientific inquiry" (404). Although published studies are by far more telling of the process of science compared with traditional school science texts, I believe that a less common type of writing by scientists may provide an even greater connection between the reader and the scientist—the first-hand account of scientific discovery, earlier dubbed the SDN. SDNs trace the scientist's journey towards the attainment of new scientific knowledge. Discovery, in this sense, excludes much of the applied sciences that generally focus on developing a practical understanding of existing scientific principles that lead to inventions like the x-ray machine. However, the meaning of discovery does not necessarily discount the creation of technology in its pursuit of new knowledge (French, under review; Hofstadter, 1985) such as Marie Curie's development of the electrometer to measure levels of radioactivity. Further, discoveries need not directly result in a major shift in paradigms (Kuhn, 1962). Thus, scientific discovery in the simplest terms is a "finding out" of some scientific phenomenon that can vary in magnitude, and the SDN highlights this "finding out" process in a way that sets it apart from most types of scientific writing.

SDNs differ from published studies that must present a series of objective, administrative steps and analyses that can be replicated and potentially refuted. The foundational importance of writing and publishing falsifiable studies (Hume, 1739; Popper, 1934) has been largely criticized over the years by a growing demand for greater transparency in the "true" nature of science as a way of knowing about our world; chance, error, intuition and creativity are important aspects in the scientific process yet often swept under the rug for the sake of presenting error-free, falsifiable evidence (Bauer, 1994; Beveridge, 1950; Feynman, 1985; Kuhn, 1962; Latour, 1987; Medawar, 1988). I do not argue against the presentation of "ideal" methods for the sake of replicability, but merely suggest the inclusion of the SDN, which allows the reader to vicariously experience both the planned and serendipitous events that led to a new understanding about the world. These less constrained accounts are often published for an audience beyond the scientists' immediate field (due to their less "ideal" description of methodology) and contain a more personal history about their work, including the motivation and emotional connections with the given topic. However, it is important to note that "less constrained" does not necessarily mean "more truthful."

As discussed in the previous chapter, narratives about the past are reconfigured accounts that reflect the narrator's perspective and her cultural values. Thus, since neither the published study nor the SDN is a completely truthful account of experience, the point of difference is the SDN inclusion of the serendipitous moments or the mishaps that occur during the phases of experimentation, which naturally lead to admissions of curiosity, confusion, frustration, excitement and sometimes despair. The following excerpts by Australian microbiologist Barry Marshall illustrate this difference:

Figure 2.1. Comparative excerpts of two accounts of a study

From, Marshall, B. (2002). The discovery that Heliobacter pylori, a spiral bacterium, caused pepetic ulcer disease. In B. Marshall (Ed.), *Heliobacter pioneers: Firsthand accounts from the scientists who discovered heliobacters, 1892-1982.* Chapter 15 (165-201).

Unbeknown to us, junior microbiology staff in the 'feces lab' where our biopsies had been relegated, were treating them much as they did feces cultures or throat swabs. Thus, expecting to see complete overgrowth of the agar plate after 48 hours, the plates were routinely examined at 48 hours and, if no unusual organisms were present, they were discarded. However, because of the workload over Easter 1982, an oversight occurred and the busy weekend microbiology technician did not examine the research plates on Saturday morning. Thus, the plates remained in the incubator until the next working day, Tuesday 13 April. On that day, the small transparent colonies of *H. pylori* were evident. The organisms were subcultured, and it was not until the following weeks that John Pearman telephoned me in an excited voice to say that the new Bacteria had been cultured. I recall he showed me some unimpressive blood agar plates and a Gram stain in which the organism certainly did not resemble the curved and spiral forms present in the histology. By then, however, they had already cultured the bacterium from another patient and were moderately confident that it was the CLO we were seeking. (p. 177).

From, Marshall, B.J. and Warren, J.R. (1984). Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *The lancet*, June 16 1984, 1311-1315.

The bacteria were S-shaped or curved gram-negative rods, 3µm x 0.5 µm, with up to 1½ wavelengths. In electron micrographs that had smooth coats and there were usually four sheathed flagella arising from one end of the cell. They grew best in a microaerophilic atmosphere at 37° C; a campylobacter gas generating kit was sufficient (Oxoid BR56). Moist chocolate or blood agar was the preferred medium. Growth was evident in 3 days as 1 mm diameter non-pigmented colonies. In artificial media the bacteria were usually larger and less curved than those seen on Gram stains of fresh tissue. They formed cocoid bodies in old cultures. The bacteria were oxidase +, catalase +, H_2S +, indole -, urease -, nitrate -, and did not ferment glucose (p. 1313).

Both excerpts above describe the study in which the growth of the bacteria called *H. pylori* was observed. The excerpt on the left comes from a chapter written by Marshall within the book that, as indicated by its title (above), deliberately relays the first-hand accounts that led to the major discovery that ulcers are generally a result of a bacterial infection and can be treated with antibiotics. The excerpt on the right comes from the scientific publication in which Marshall and his colleague Warren reported their initial findings on *H. pylori*. The highlighted portions in both excerpts have a similar propositional quality in that the Gram stain containing the cultured bacteria is described as being less curved in shape. However, the surrounding context for each description is quite different. Within the book excerpt, Marshall informally comments on the accounts that led to the first observation of the elusive bacteria *H. pylori*. The mishaps and idiosyncratic details that happen during experimental procedures find a place within this piece by Marshall, unlike the excerpt from the published study on the right. Marshall describes in the book excerpt, for example, that an oversight on the part of a technician led to the further incubation and a serendipitous observance of bacterial growth. This piece of the "real story" finds no place in the published study showcased to the right. Instead, Marshall and Warren report on the most advantageous conditions (temperature, medium, etc.) for culturing the bacteria, all such information realized after the mishaps and delays.

One of the most important purposes of a published study is the exposure for potential replication and subsequent validation testing by other scientists in the field (Kuhn, 1962). Thus, the idiosyncratic details are not useful in this context; a clearly defined, objective set of procedures and conditions is required for replication and possible refutation.

These contrasting styles of reporting reflect the contrast between the general purpose for documenting the accounts leading to a discovery and the purpose of reporting replicable findings. In the first, a scientist intends to tell her story as a moment in history to be recorded and communicated to a wider audience, whereas in the second, a scientist wants to challenge others within her field to scrutinize the validity of her findings and try to find fault with them. Both purposes have value, albeit in different settings. First-hand accounts of discovery as represented by Marshall's (2002) excerpt about the importance of the mistake, in my opinion, would be of particular value for science education on two levels: a) exposure to the "real" process of science and b) engagement in the unique human qualities of scientific experimentation or exploration.

First, as mentioned earlier, students and teachers alike have little understanding of the nature of real scientific processes that often include repeated failures and chance encounters. The commitment of scientists to stay the course in the face of repeated failures and the key roles of chance and serendipitous events are largely neglected in science curricula. Science is thus viewed as a step-by-step process rather than what it is in reality—a creative endeavor that often involves unexpected events and "spur-of-the-moment" decision making.

A second reason for highlighting the mistakes made during the science process is the potential engagement for students and teachers alike. The unexpected events of chance and mishap make bring a human quality to the endeavors of science. Sometimes scientists can be wrong, and sometimes communities can be wrong about scientists. Learning about this aspect of the scientific process through the discovery narrative may engage readers in the subject matter in a way that traditional texts do not. For example, Marshall's account of an oversight in the culturing process may invite students to identify with the human qualities of the scientists in ways that the more traditionally scientific accounts would not. Thus, I turn my attention to first-hand accounts of discovery, which provide a more complete picture of the "messiness" involved in scientific endeavors. I will refer to these first-hand accounts as scientific discovery narratives (SDNs) throughout the rest of this discussion.

The informal, conversational style of Marshall's SDN of *H. pylori* is a clear contrast to the more formal, objective style of the published report by Marshall and Warren. However, formality of discourse used in a text is not a consistent differentiating factor between these two forms of writing; formal published reports were not always devoid of less formal, personal reflection. The idiosyncrasies of personal style of presentation as well as intended audience of readers can blur the distinction between these two genres of primary science texts. Michael Faraday (1859), for example, commonly wrote in a very personal style even in his formal treatise, an official publication from The Royal Institute of London. Faraday reports on his findings without an explicit account of his discoveries, yet commonly mixing personal observations about his experiences with philosophical observations about the principles involved in the scientific process:

> It is right that we should stand by and act on our principle; but not right to hold them in obstinate blindness, or retain them when proved to be erroneous. I remember the time when I believed a spark was produced between voltaic metals as they approached to contact (and the reasons why it might be possible yet remain); but others doubted the fact and denied the proofs, and on re-examination I found reason to admit their corrections were well-founded . . . . You can have no idea how often and how much, under such an impression, I have desired that the marvelous descriptions which have reached me might prove, in some points, correct; and how frequently I have submitted myself to hot fires, to friction with magnets, to the passes of hands, etc., lest I should be shutting out discovery; --encouraging the strong desire that something might be true, and that I might aid in the development of a new force in nature (474-475).

Faraday's open admission is not a SDN. The purpose for including this example is to further clarify the distinct characteristics of SDNs. Faraday reveals the emotional twists that scientists often experience in their work, but he merely touches the notion of discovery and does not offer details about his work with voltaic metals. He uses the issue of discovery as an example of the importance of having critical friends, no matter how depressing being wrong can feel. Such personal expressions are generally absent from scientific manuscripts published after the 19th century once Comte's argument for positivism (i.e., that "true knowledge" can only be attained by what is observed by the senses) started to take root in scientific discourse (Giddens, 1974; Mises, 1951). SDNs, on the other hand, seem to sidestep the modern rules for scientific publications since they are not necessarily intended to be replicable and objective. Thus, accounts of discovery vary greater in the level of formality and intimacy, depending on the intended reading audience. The following excerpts by Nobel-prize winning molecular biologist Francis Crick (1988) and famous botanist and inventor George Washington Carver (1911) illustrate this variety. Crick offers the reader with a lecture-style overview of evolutionary theory that was the foundational stepping-stone for his work on DNA. He uses a formal, sometimes poetic style, but minimizes his use of highly technical terms as he is writing for the general, literate audience:

> When [*The Origin of Species*] was published in 1859, it immediately ran through several reprintings and did indeed produce a sensation . . . It needed only the discovery of genetics, originally made by Gregor Mendel in the 1860s, and, in this century, of the molecular basis of genetics, for the secret to stand before us in all its naked glory" (p.25).

Carver, on the other hand, uses a more intimate, conversational style in his letter to his friend, Booker T. Washington, about discovering a more affordable substance for making house paint. There is no intended lecture in his letter and it is clear that Carver had been discussing his work with clay in previous exchanges with Washington:

> My dear Mr. Washington:
> I think you will be interested in the following observations:
> I went out into the country Saturday to make some observations with reference to my work with the clay, and stopped at Mrs. Pugh's. I found that she had a house with four rooms, all of which were whitewashed with the clay, being perfectly white. She tells me that she has used it for years, and has not used a bit of lime. She had some of the clay mixed up in a bucket when I got there . . ." (p. 113)

These excerpts by Crick and Carver represent different levels of formality and shared understanding due to their intended audience. Further, there is a great difference in the magnitude of the discoveries described. Whereas Crick's book presents groundbreaking evidence for the genetic structure of life, Carver's letter presents evidence for a new, cost effective material for making house paint. Regardless of these differences, the common element in both pieces is the discovery and the telling of how that discovery came to light. The focus of my investigation is to determine if there are common components in the telling of these discoveries, regardless of the obvious diversity in terms of intended audience, formality, shared understanding and magnitude of discovery. The critical first step in this investigation is the collection of a representative sample of SDNs.

Analytic methods for comparing SDNs

This section addresses the criteria and analytic methods used in the process of identifying common elements of SDNs. I explain the process of gathering, outlining and mapping each of the SDNs in my sample. Based on a "macro structural" approach from Kintsch and van Dijk (1983), I found the four general ideas that were mentioned in the beginning of this chapter—love for the work, the known, the unknown and the exploration process—each of which is discussed in this section.

Selection criteria

The collection of SDNs for this analysis was the result of four selection criteria. First, as previously stated, the present analysis is focused on accounts that explicate the process of how new scientific knowledge came to light, which may include the creation of various applications in the pursuit of that new knowledge. Second, I limited my search to first-hand written, published accounts as this helped to simplify the sampling process and to avoid any dictated speeches or lectures that may have a varied set of objectives other than to explicate the discovery process. Biographies like Mahon's (2003) account of James Clerk Maxwell's discovery of electromagnetic force were discarded in light of the variability in the kinds of information presented in them. Depending on the interests and expertise of the author, the biographical narrative took on more than just the discovery of a scientific concept. Marrin (2002), for example, shares a great deal of the historical background of smallpox and how it was characterized through time as he tells about how physician Edward Jenner during the late 1700's came to his discovery that cow pox can help vaccinate people from the deadly virus. Isaacson (2007) takes a more socio-cultural view of Albert Einstein's work and how issues of anti-Semitism and frequent philandering affected his work. These stylistic differences are a bit problematic since my goal is to determine a consistent pattern of components that exist across all accounts of scientific discovery. Further, popular histories like these have been criticized for dramatically inflating the discoveries by over emphasizing and often times ignoring or misattributing important contributions from other scientists (Allchin, 2003; Westerlund & Fairbanks, 2004). Thus, an important criterion was selecting only first-hand accounts in hopes of at least limiting such biased writing to the scientists themselves.

A third decisive factor in the SDN sampling process was to choose written accounts in which the author explicitly communicates her intention to write an account of discovery. These accounts would presumably present a more complete account of how the scientist came to know something about a particular phenomenon, which may include serendipitous events as exampled above in Marshall's SDN excerpt. Finally, it was important that the collected sample of SDNs represented the diverse population of scientists who have written such accounts. Thus, I made every effort to include SDNs written in different time periods by both male and female scientists from various countries around the world.

It is important to note two unavoidable limitations in my analysis. First, all selected works had to be analyzed in English (my dominant language), thus a full account of the stylistic differences among SDNs, as exampled by the Crick/Carver comparison above, would be difficult due to the overlay of editing and translation of almost half of the SDNs in my sample. However, a more rhetorical analysis of the different stylistic methods used in composing these SDNs (e.g., use of particular phrases or words for describing the discovery process) is unnecessary. My scope limited to "what" information is included in these accounts rather than "how" these components of information were written, which is congruent with the goal of my study. Thus, the sole focus on English texts in determining common components of information across SDNs is not a significant issue for this analysis.

Another unavoidable and more unfortunate limitation is that as a result of my sampling from SDNs written or translated into English, I had inadvertently limited my analysis to accounts by scientists within the western hemisphere. Written work by scientists within the middle or far eastern regions of the world are typically not written or translated into English and are largely inaccessible to western researchers. The only non-western work I was able to find is the translated works of Jabir, the 8th century Middle Eastern philosopher and chemist who painstakingly wrote accounts of his experiments with naturally occurring metals (translated and edited by Holmyard & Russell, 1997). Jabir's work was written during a time when science was deemed more as a magical art; many scholars during his lifetime were in search of an "elixir" with the believed powers to purify metals such as tin, lead and copper and transform them into silver or gold. The inclusion of Jabir's account contributes to the diversity within my sample, both in terms of culture and time period, but there would have been greater balance in this sample of SDNs if I were able to include a modern account from this region as well. Thus, I must admit to the possibility that my sample of SDNs may not fairly represent all the existing SDNs in the world.

A representative sample of SDNs

After a full year of looking for potential candidate SDNs according to the selection criteria described above, the final sample includes the twelve SDNs presented on the following page.

Table 2.1. Descriptive List of First-Hand SDNs

Scientist	Source	Subject	Noted Discovery
Gerber, also known as Jabir (Iraq)	Jabir (cir. 8th cent.). In E.J. Holmyard and R. Russell (Eds.), *The works of Gerber*. West Glacier, MT: Kessinger: 1928.	Chemistry (Alchemy)	Questioned and qualified the essential elements of metals (cir. 8th century)
Galileo Galilei (Italy)	Galilei, G. (1610). *Starry messenger* (*sidereus nuncius)*. translated by A. van Helden, Chicago: 1989.	Astronomy	First to analyze and document the topography of the moon; discovered the moons of Jupiter (1608-1609)
Edward Jenner (UK)	Jenner, E. (1798). An inquiry into the causes and effects of the variolæ vaccinæ, or cow-pox. In C.W. Eliot (Ed.), *The Harvard classics*. NY: P.F. Collier & Son, 1909-14. Bartleby.com, 2001. www.bartleby.com/38/4/. [Date of Printout].	Medicine	Discovered a link between cowpox and smallpox, leading to his development of a small pox vaccine (1796)
Juan Carlos Finlay y Barrés (Cuba)	Finlay y Barrés, C.E. (1881). Original correspondence. In M.C. Khan (Ed.), *Carlos Finlay and yellow fever*. NY: Oxford University Press: 1940.	Medicine	Discovered the link between the mosquito and yellow fever (1881)
Marie Curie (Poland & France)	Curie, M. (1904). *Radioactive substances*. London: Chemical News Office.	Chemistry	Co-discoverer of elements radium and polonium (1902)
Kristian Birkeland (Norway)	Birkeland, K. (1913). The Norwegian Aurora Polaris Expedition 1902-1903: Vol. I, On the cause of magnetic storms and the origin of terrestrial magnetism (Unknown Binding).http://ia340919.us.archive.org/2/items/norwegianaurorap01chririch/norwegianaurorap01chririch.pdf	Physics	Correctly deduced the solar-magnetic origin of the Aurora Borealis (1902-1903)
George Washington Carver (USA)	Carver, G.W. (1911). Personal correspondence with Booker T. Washington. In J.R. Kremer (1987), *George Washington Carver: In his own words*. Columbia, MS: University of Missouri Press.	Agriculture (botany) and Chemistry	Adapted the use of clay and various plants for making different colors of house paint (1911)
Jane Goodall (UK)	Goodall, J. (1971). *In the shadow of man*. NY: Houghton Mifflin.	Biology	First to document group behaviors of chimpanzees (1960)
Francis Crick (UK)	Crick, F. (1988). *What mad pursuit: A personal view of scientific discovery*. NY: Basic Books.	Molecular Biology	Co-discoverer of the structure of the DNA molecule (1953)
Barry Marshall (Australia)	Marshall, B. (2002). The discovery that Heliobacter pylori, a spiral bacterium, caused pepetic ulcer disease. In B. Marshall (Ed.), *Heliobacter pioneers: Firsthand accounts from the scientists who discovered heliobacters, 1892-1982.* Chapter 15 (165-201). MA: Blackwell Publishing.	Microbiology (Medicine)	Co-discoverer of the H. pylori's role in gastritis and peptic ulcer disease (1982)
Alan Guth (USA)	Guth, A. (1997). The inflationary universe: The quest for a new theory of cosmic origins. Reading, MA: Perseus Books.	Theoretical Physics	Conceptualized "cosmic inflation" (1979)
Eric Kandel (USA, Austrian-born)	Kandel, E.R. (2007). *In search of memory - The emergence of a new science of mind*. NY: WW Norton & Company.	Neuroscience	Co-discovered the synaptic origins of short and long-term memory (1972-1983)

As mentioned earlier, this is not intended to be an exhaustive list but only a representative sample of scientists who have documented their discoveries in SDNs. Further, it is not clear when documented accounts of scientific discovery began to appear; the earliest record I have found is chemist Jabir's account from the 8th century (later translated by Holmyard and Russell in 1928). Earlier descriptions of discovery may exist; for example, there may be accounts of discovery in the numerous manuscripts by the Greek physician and philosopher Galen, who lived during the first century, but his works are largely inaccessible. A little more than half of the collected works in this sample represent English-speaking countries (Australia, UK and USA), a consequence of selecting from accounts that were available in English, as described earlier. However, since my intention for this analysis is to develop adapted versions of SDNs for middle school students in the United States, the English-dominant accounts would not be inappropriate for the present study.

Framework for Analysis

Because the sampled SDNs vary greatly in length, I chose an analytic approach that focused on global-level meaning units, or big ideas of a text. The macro-structural approach to narrative analysis by Kintsch and van Dijk (1983) seemed the most appropriate guiding principle for analyzing and comparing SDNs within this sample. As mentioned in the previous chapter, these researchers developed a structural account of narrative similar to the conflict narrative structure by Labov and Waletzky (1967). Kintsch and van Dijk's (1983) approach to analyzing narratives is of particular interest in that the focus is on the big ideas (macro propositions), which are formed from the building up of smaller ideas (micro propositions) within a given text. These macro propositions could be viewed as "summarizing points" for each paragraph, section or chapter within an SDN, depending on the length. The longer the SDN, the more subtitles and headings were present, which helped in identifying the macro structures within a given text. The following introductory paragraph to Curie's (1904) *Radioactive Substances* illustrates this process of identifying a macro structure:

> The object of the present work is the publication of researches which I have been carrying on for more than four years on radio-active bodies. I began these researches by a study of the phosphorescence of uranium, discovered by M. Becquerel. The results to which I was led by this work promised to afford so interesting a field that M. Curie put aside the work on which he was engaged, and joined me, our object being the extraction of new radio-active substances and the further study of their properties (p. 3).

The sentences within this initial paragraph essential address why Curie decided to study the curious glow from uranium rock. The large idea, or macro proposition, here would be that for Curie and her husband, studying radioactive bodies is really interesting. The following paragraph introduces a new point:

> Since the commencement of our research we thought it well to hand over specimens of the substances, discovered and prepared by ourselves, to certain physicists, in the first place to M. Becquerel, to whom is due the discovery of the uranium rays. In this way we ourselves facilitated the research by others besides ourselves on the new radio-active bodies (p.3).

Curie follows her initial statement about her own interest in radioactivity with the idea that other scientists should be included in this important area of scientific research. Thus, two major ideas, or macro propositions, are present within the first page of Curie's book. I read and analyzed sections of each sample SDN in a similar fashion, paying particular attention to the first chapters of books since many of these introductory sections contained the macro propositions that are discussed in greater detail in subsequent chapters. Throughout the process of outlining the macro structure of all SDNs, I soon became aware of a pattern in the types of ideas expressed in these accounts.

Common Elements of the SDN

The following four common components surfaced as a result of outlining the macro structure of each SDN: a) love for the work, b) acknowledgement of what is known, c) identification of the unknown and d) the exploratory or experimental process that led to the discovery. In the following discussion, I describe each of these components respectively, and include illustrative excerpts from my sample of SDNs. Although there seems to be some overlap in how these elements are presented (especially in regards to expressions of love for the work), the order suggested above generally mirrors the order of big ideas emphasized in each SDN.

For the Love of the Work

Clearly evident in all the SDNs within my sample was how much the scientists loved their subject matter and the current research problem that had caught their attention. This love was not expressed with flowery sentiments or cheery, affectionate praise, rather, it seemed more like a deeply rooted marriage that has yet to lose its passion. In many of the accounts are bouts of frustration and despair, as illustrated by Goodall's (1971) response to her repeated failures in trying to find her primates:

> Later I realized how lucky I had been during the fruiting of the msulula tree. I probably learned more during those ten days than I did during the eight depressing weeks that followed. Search as we would, we found no other large fruiting tree or group of fruiting trees. We went up most of the twelve valleys of the reserve, but the undergrowth was often very thick, and although the sound of the streams drowned any noise we made it also effectively camouflaged the sounds that might have warned us of the whereabouts of chimps. Those we did see were usually so close by the time we came upon them that they fled instantly. I now can well imagine how many times they must have seen us coming and silently vanished without our ever being aware of their presence (p. 21).

Yet it seems that love for her work allowed her to persevere, even in the face of obstacles such as the state of isolation:

> As the weeks passed, however, I accepted aloneness as a way of life and I was no longer lonely. I was utterly absorbed in the work, fascinated by the chimps, too busy in the evenings to brood (p. 50).

Such devotion was also evident in Birkeland's (1913) expedition for discovering the underlying causes for what we now call the aurora borealis. Birkeland spent seven years and all of his own resources in his quest to find out what caused the northern lights, and in particular whether the sun's rays have any role in this phenomenon. Throughout his descriptions of horrible, even near-death experiences in reaching remote destinations, he maintains an exuberant tone:

> At last we started, each in our pulk, after the guide had solemnly asked us if it were really our intention to try to get back to Gargia in this weather. We could not see more than a few yards in front of us, but we were quite determined to try. The couple of hours spent in the descent were the most exciting I have ever gone through. It was now that our guide showed himself to be the adept that I had been told he was. It was wonderful to see the way he ran to the right or to the left, to find tracks or take a course, and how he drilled the reindeer when they became unmanageable . . . . We then at last got something to eat, not having tasted food all through the terrible journey; and then we once more turned our attention to Helland-Hansen's hands, which were in a terrible state, and dressed them as well as in the mean time we were able. And in spite of everything, our spirits now rose high, in our intense delight at having at any rate not lost our lives (p. 3)

Love for the subject matter seems to begin with some form of attentive curiosity and develops into a devotion for learning as much as possible about this subject, with a heightened focus on the current research question. As illustrated earlier, the initial spark for Curie (1904) was Henri Becquerel's discovery of uranium ore and its mysterious glowing quality. For Cuban physician Finlay y Barrés (1940), his focused interest about the *Culex* mosquito began with the development of a new idea about the origins of yellow fever: "So far as I can remember, what first suggested to my mind the idea that the *Culex* mosquito might be the intermediary host of the yellow fever germ was an account published in Van Tieghem's Botanique of the life-cycle of the *Puccinia graminis,* in which I was very much interested" (p. 73).

The scientist may also identify the beginnings of interest early in his or her work, before the focused interest on the particular topic has been fully shaped. Neuroscientist Kandel (2007), for example, attributes his focus on the neuron to the guidance of a mentor in medical school who argued that in order to gain a better understanding about how the brain works, it should be studied at the cellular level. This focused attention can develop into a heated interest, as described by Carver (1911) and his new finding about clay: "It really grows more interesting the more I investigate it. I shall make further investigations with reference to the deposits on our own land just as rapidly as I can get to it" (p. 113). Carver continues to gleefully describe scouring the countryside for various plants and other potential paint materials, with no criticism about the time and distance required to retrieve such items.

At times the scientists' language can be effusive, even poetic, in their expression of love for their work, like Galileo's (1610) gushing jubilation about seeing the topographic features of the moon: "It is a very beautiful thing, and most gratifying to the sight, to behold the body of the moon, distant from us almost sixty earthly radii, as if it were no farther away than two such measures . . ." (pp 27-28). Even though this "gratifying sight" required meticulous movements to capture various parts of the moon (he was unable to see the entire moon in his field of view) over many nights, he gives not a single complaint. Thus, the scientists' enthusiasm for their work overshadows any imaginable obstacles that naturally arise in the process of discovery.

These general expressions of love by the scientists stem from different aspects of their work. Whereas Goodall (1971), Curie (1904) and Galileo (1610) seemed more enthralled with the central concepts of the science, others are more focused on resulting effects of their discoveries. Finlay y Barrés (1940), for example, was spurred onward by the desire to know the causes and potential cure for yellow fever. Similarly, physician Jenner (1798) focuses on the numerous cases of people infected by cowpox during his time, building a theory that would eventually lead to a vaccine for the more deadly virus, smallpox. Thus, it seems that for the two practiced physicians in my SDN sample, love for the work largely stems from a love for people and finding ways to save lives.

The love expressed by the chemistry pioneer Jabir (1928, cir. 8th century), seems to stem from the "art" of conducting science in which there is a sense of a spiritual devotion. Jabir begins his treatise with the following statement: "We with continued and frequent diligence of Labour, and great Study equivalent, not without most profound and serious thoughts, &c. expose publickly to your view, the Investigation of this most noble Science, that the subsequent Volumns may the better and more clearly be understood, searched into, and found, they may the more easily and readily be brought to effect" (p.3). Throughout Jabir's treatise about the nature of various metals, there is a deeply ingrained reverence for the act of systematic investigation as though it were a holy, or reverently magical practice. During this time period, the practice of alchemy dominated much of the Middle East. Alchemy is founded on the premise that certain metals can be "transmuted" into other types of metals, like turning tin or lead into gold or silver. According to literary scholar Holmyard (1990), this transmutation was thought to occur with the application of a substance called either the "Elixer" or "Philosopher's Stone," the attainment of which was a central issue for alchemists. The potential for this transmutation elevated the status of alchemists to a spiritual, theistic level in that the study of metals was thought to be a practice that would bring the scientist closer to god.

Regardless of its specific origins, it seems that love for the work plays an important role in motivating all of the scientists in this sample to continue onward in the face of mishaps or temporary failure. Repeated trials and long waiting periods can hardly be viewed as enjoyable, yet such experiences seem tolerable in light of the larger goal of learning something new and contributing to the ever-growing body of knowledge in a particular field.

Acknowledging What is Known

In his book, *Thematic Origins of Scientific Thought*, physicist and scholar Holton (1988) argues that no scientist begins their work with a blank slate. Even the ancient alchemist Jabir (1928, cir. 8th century) referred to what other "artists" (as scientists were called during this time) had discovered: "We find Modern Artists to describe to us one only *Stone*, both for the *White* and for the *Red*; which we grant to be true" (p.6). No matter the level of detail, all scientists within the SDN sample consistently referred to existing knowledge about particular topic of focus. Some of the accounts contained a vague mention of other individuals, like Galileo's (1610) *Starry Messenger*: "About ten months ago a report reached my ears that a certain Fleming had constructed a spyglass by means of which visible objects, though very distant from the eye of the observer, were distinctly seen as if nearby" (pp. 28-29). Jenner (1798) offers a little more deference in his reference to the work of others:

> I have lately also been favoured with a letter from a gentleman of great respectability (Dr. Ingenhousz), informing me that, on making an inquiry into the subject in the county of Wilts, he discovered that a farmer near Calne had been infected with the smallpox after having had the cow-pox, and that the disease in each instance was so strongly characterized as to render the facts incontrovertible. The cow-pox, it seems, from the doctor's information, was communicated to the farmer from his cows at the time that they gave out *an offensive stench from their udders.* (from "II. Further Observations on the Variolæ Vaccinæ, or Cow-Pox. 1799," para. 2)

General reference to "the known" can also be attributed to an entire field, like molecular biologist Crick's (1988) description about what was known about genetics before the breakthrough on DNA: "Genetics tells us that, roughly speaking, we get half of all our genes from our mother, in the egg, and the other half from our father, in the sperm" (p. 31). Neuroscientist Kandel (2007), on the other hand, includes all of the key scientists involved in developing the field of neuroscience and how many individuals from various fields provided a foundation from which he was able to confirm the synaptic nature of the neural network. Similarly, physicist and cosmologist Guth (1971) describes in great detail the key members and institutions involved in sorting out the origins of our universe.

Even when very little is known, as the case of Goodall's (1971) chimpanzees and Birkeland's (1913) northern lights, there is still mention of foundational knowledge. Both of these scientists needed to be well prepared to work in their respective environments and thus described the guidance and support of existing knowledge to conduct their explorations into the unknown.

All SDN descriptions about "the known" seem to diverge from the canonical background section that is often present in published scientific studies in the last fifty years. Specifically, there are two general distinctions between SDNs and modern published studies with regards to "the known"—a) the scope and focus of existing knwoledge and b) the presented order of the problem (unknown) in relation to the existing knowledge. Without the urgency to support an argument with well-known references, it seems that the scientists could focus more on the parts of existing knowledge that helped in fostering their own thinking about a particular problem, such as exampled by Jenner (1798) above is exemplified. Guth (1997) provides some insight on this point:

> [This book] is definitely not intended as a thorough history of the research relevant to this work . . . I have attempted, nonetheless, to honestly describe the roles of the people with whom I interacted, since these were the players in the drama that I wanted to recreate . . . I have completely omitted a number of related lines of research . . . The omission of these contributions is by no means intended as a judgment of their importance, but merely a consequence of the personal viewpoint that unifies the book" (pp. xiv-xv).

In a published study, existing knowledge is explicated to the extent to which it supports the primary argument of a research piece. By contrast, authors of SDNs incorporate existing knowledge to further contextualize the journey towards discovery. This is not to say that SDNs are not meant to be compelling accounts about the importance of a new discovery; certainly Galileo (1610) had much to do to convince his patrons who were under the powerful influence of the Vatican about the possibility that Copernicus was right in his theory. For SDNs, it seems that the compelling argument *is* the story of discovery, including the key individuals that played a direct role in this discovery. Thus, in SDNs, the existing knowledge leads the discoverer to identify the problem, rather than first presenting a problem and then offering sufficient evidence for the argument, which is generally the case for modern published studies.

Identifying the Unknown

According to renowned mathematician Polya (1945), before an individual can learn something new, one must have a full understanding of the problem itself. All SDNs highlight the unknown and just as with the case of the previous section, there is variability in the amount of detail in and the motivating purpose of the unknown issue. More specifically, some SDNs focus on an unknown of urgent practical importance while others simply explore the unknown for the sake of knowledge acquisition without a particular practical agenda. Curie's (1904) work on the mysterious glowing nature of the uranium ore, for example, was motivated by her curiosity in the elements and had no immediate connection with practical applications of this knowledge. Her discovery of two radioactive elements (radium and polonium) was a significant contribution to the field of chemistry, but the enormous impact that the uranium ore problem had on saving (and destroying) human lives was not realized until much later after her death. Similarly, Crick (1988) focused on an unsolved problem in biology, the molecular nature of living organisms, which eventually led to the discovery of DNA and a new field of genetic engineering. Crick writes:

> And now we can approach the baffling problem that appeared to face us. If genes are made of protein, it seemed likely that each gene had to have a special three-dimensional, somewhat compact structure. Now, a vital property of a gene was that it could be copied for generation after generation, with only occasional mistakes. What we were trying to guess was the general nature of this copying mechanism (p.35).

The discovery by Crick and his colleagues was far more than just knowledge acquisition for the field of biology. This new understanding about the fundamental structure of life has had a significant impact on both the fields of medicine and agriculture, leading to wide spread practices of genetic testing and modification for a variety of purposes.

For some pioneers of scientific discovery, the problem is an urgent issue from the very beginning. The root cause of yellow fever, for example, was a mystery that had confounded many physicians like Finlay y Barrés (1881) who endeavored to find the source of this disease that had claimed countless lives. Marshall's (2002) problem with the unknown cause gastric ulcers also had an immediate impact on human lives: "Robin and I were acutely aware that this was merely an anecdotal observation and so we were biased against observing any clinical response in this patient. The patient himself, however, was ecstatic in his discovery that an antibiotic treatment was able to completely eliminate his extremely severe and chronic gastric symptoms" (p. 173). Discovering the source of the extreme pain that people experience with ulcers clearly improved the quality of life for so many around the world.

Not all endeavors in science have recognizable or immediate utility. Guth (1997), for example, focused his research on the long debated issue known as the "horizon problem," which concerns the tension between the classic "big bang" theory and more recent data on background radiation. Although Guth's eventual "inflationary" theory did not save lives, it moved the field of cosmology closer to understanding the origins of our universe.

As described earlier, alchemists like Jabir (1928, cir. 8th century) were focused on finding the *Elixer,* which led to Jabir's specific problem of understanding the conditions and substances that would *corrupt* (i.e., weaken) or *perfect* (strengthen) metals. Although he never found or created a substance that could turn lead into gold, Jabir's focus on the unknown elemental structure of metals was an inadvertent contribution to a further distancing between science and the practice of alchemy and a step toward the development of the field of chemistry.

Sometimes the unknown posed by the scientist is a direct confrontation to conventional thinking. Galileo (1610), for example, was skeptical about the idea that all of the heavenly bodies in space revolved around the earth and suspected that Copernicus was right about his argument that the sun was the center (Finnocciaro, 1997); thus he jumped on the opportunity to develop a telescope powerful enough to find the evidence needed to support the Copernican view.

Oftentimes new knowledge is needed to settle inconclusive findings in a field, as illustrated by Birkeland's (1913) quandary: "By examining the literature on the subject, I soon saw that there were no definite results that could decide the question, although our idea of the constitution of matter presupposes that the atoms in a non-electric piece of metal are positively charged, and that there are corresponding free negative electrons between the molecules" (p. 716). Similarly, competing theories in the neural mechanisms of the brain intensified the issue for Kandel (2007) to find a way to learn more about the fundamental elements of neurons.

The "unknown" in SDNs may be spurred on by an urgent call from the field or the masses as evidenced in the examples described above. However, even a modest discovery such as Carver's (1911) clay-based paint can be worth writing about, depending on the purpose and prerogative of the scientist. Published studies, by contrast, fulfill a prescribed set of expectations, many of which are largely determined by the peers of the author rather than the authors themselves. Further comments on this distinction will be saved for the concluding remarks of this chapter.

Explorations and Experiments

There are many paths to scientific progress, and if our only lens through which to view this progress is the published scientific study, we would fail to appreciate the great diversity of the "culture of learning" that is science. SDN descriptions of the actual experiments or observations leading to discovery vary in nature of study and detail depending on the length and degree of formality of the account, and because of this variability, we are able to see the multiple ways in which learning can take place.

Some accounts seemed to be written for the reader to follow as a guide, with even greater detail than the methods section of a typical published scientific study. The lengthiest account within my sample comes from Birkland (1913), who in a thousand-page volume presents both the mathematical investigations and the actual experimental procedures with numerous meteorological diagrams, graphs and pictures in order to "make clear the movement of electrical corpuscles from the sun to the earth, which have been carried out, with a perseverance and ingenuity worthy of all admiration" (p. iv). Even the exact dimensions and materials used in building the twenty-five observatories were included in this expedition, down to the number of nails used in their construction, revealing potentially helpful details for anyone who wishes to follow in his footsteps. Although Galileo's (1610) *Starry Messenger* is significantly shorter in length and thus does not fully describe each step in his investigations, a reader is nonetheless able to follow the process that leads to his first findings about the dark "spots" on the moon:

> [T]hese I shall call the "large" or "ancient" spots, distinguishing them from others that are smaller in size but so numerous as to occur all over the lunar surface, and especially the lighter part. The latter spots had never been seen by anyone before me. From observations of these spots repeated many times I have been led to the opinion and conviction that the surface of the moon is not smooth, uniform, and precisely spherical as a great number of philosophers believe it . . . (p. 31).

The relative brevity and informality of Carver's (1911) letter understandably affords little mention of the exact processes in the investigation of clay-based paint and ways to add different colors, such as making "a beautiful black color by simply taking some of the clay and mixing a small amount of boiler black with it," which was essentially the soot extracted from boilers (p. 114). In fact, Carver's letter may not seem very "scientific" in that he does not pose any formal hypotheses in his description, nor does he seem to implement any formal procedures for his investigation. Carver's account of the clay is similar to the kind of everyday narrative that was discussed in the previous chapter. However, I believe that it is accounts like these that best reveal the inherent "humanness" of science. All of us are a part of a culture of learning, and yet the culture of learning in science is generally viewed as an elite practice that has little to no immediate relevance in everyday life (Latour, 1987; Lemke, 1990). The SDNs represented here reveal the relevancy of discovery in the familiar context of learning something new.

Students in science classrooms are often instructed to follow the "scientific method" when conducting experiments, which generally includes (with great variability) the following set of procedures: a) ask a question, b) gather information and resources, c) form a hypothesis, d) test the hypothesis, e) analyze data, f) draw conclusions, and g) report results. This set of procedures is usually displayed on large posters in the science classroom, a reminder to all about the "proper" way of conducting experiments. There are variations in the configuration and labeling of these steps, as illustrated by the following examples:

Figure 2.2. Various poster images on the scientific method

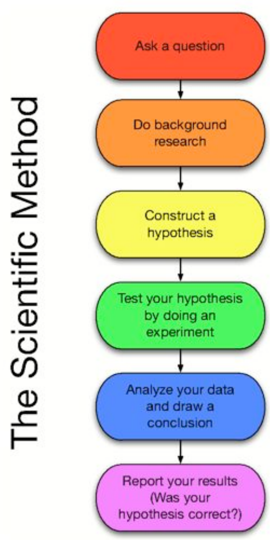
a)

www.sciencebuddies.org/

b)

www.mghs.sa.edu.au/.../ workingScientifically.htm

c)



kbagdanov.wordpress.com

d)

web2.newtown-h.schools.nsw.edu.au/ Science

Although the posters labeled "a" and to some extent "d" encourage a revision in thinking about a previous hypothesis, there does not seem to be a place to be completely wrong about an idea. In other words, it is unclear whether students have a solid understanding about what it means to revise a hypothesis. The scientific method has long been criticized for its myopic view of science; the steps oversimplify the complex and dynamic process of science and the conceptually oriented models that are employed by scientists across different fields (Abd-El-Khalick, et al, 2004; Cartwright, 1983; Giere, 1988; Nersessian, 1999). A recent movement to enhance this traditional model has led to a greater focus on the term "science inquiry," which is defined as a process in which students engage in observations and make inferences while employing a variety of methods for investigations and experiments (National Science Education Standards, 1996). The purpose of this shift in emphasis is to give students the opportunity to engage in science as a scientist actually would; the more authentic this practice is, the more a student will understand about the targeted concepts as well as the practice (Gopnik, 1996).

Research in science education is demanding that schools adopt a more complex dynamic view of science as a way of knowing about the world, but there seem to be no explicit guidelines about which complexities to reveal to students. Should we explain, for example, how scientists can start completely over in their work? Consider the following excerpt by Finlay y Barrés (1881) who had been convinced that the cause of yellow fever had to do with the presence of "excessive alkalinity" in the atmosphere, and so performed a series of experiments, only to report the following:

> In consequence thereof, I feel convinced that any theory which attributes the origin and propagation of yellow fever to atmospheric influences . . . as utterly indefensible. I have, therefore, been obliged to abandon my former ideas and shall now endeavor to justify this change in my opinions . . . (p. 65).

People engaged in a culture of learning, including scientists, experience countless revisions and complete overhauls of their conceptions about the world. And what about the essential elements of "hoping and waiting" for many scientists? Goodall's (1971) explorative process involved a great deal of waiting and hoping to catch a glimpse of chimpanzees in action:

> I was waiting on one side of a narrow ravine, hoping that chimpanzees would arrive in a fruit-laden tree on the opposite slope. When I heard the deliberate footsteps of approaching chimpanzees behind me I lay down flat and kept still—for often, if the apes saw me on their way to feed, they changed their minds and fed elsewhere (p. 58).

Inadvertent successes, such as described by Marshall (2002) earlier can be a crucial factor in discovery (Holton, 1973) and seem to be largely unexplored in science education. The use of SDNs like those sampled for this analysis could be useful in illustrating the complex and dynamic nature of discovery for students.

Conclusions

SDNs are narratives, like published scientific studies, that describe a problem within a given field and the process that led to a solution, which is the discovery. In this chapter, I have examined the qualities that separate SDNs from other scientific writing through the lens of the four common components—the love for the work, the known, the unknown and the process towards discovery. SDNs vary greatly in detail and emphasis across these four common components and often include incidents involving mishaps and chance as illustrated in the examples above. With such variability, one might wonder what the value is in studying such accounts. Would it not be more appropriate to focus on published studies, which are more structurally uniform and are considered as the "standard form" of scientific writing? I argued in the first chapter that studying scientific narratives in the form of published studies provides a valuable lens for discussing many aspects of scientific work that are currently ignored in the classroom. However, I have made the additional argument in this present chapter that it is equally important to consider other scientific narratives, like SDNs, that help students understand the great variability in the culture of scientific learning. For my study, I plan to focus on the SDN and investigate whether this type of scientific narrative has a unique effect on the interest, understanding and memory recall of middle school students.

Identifying the common elements in SDNs is only the first step to adapting and developing texts that will be appropriate for middle school students, which is the central purpose of my study. The four common components described above provide a guide for ensuring that the experimental SDNs are comparable to the first-hand accounts yet written at a level appropriate for young readers with little to no prior knowledge of the targeted scientific information. Thus, the four-component foundation for writing such narratives is only part of what would make a text worth reading; it is equally important to consider issues of readability, coherence, engagement, etc. In the following chapter, I describe the process of developing four engaging, accessible and cohesive science texts for middle school readers. These four texts differ in terms of genre (SDN and non-SDN versions) and topic (radioactivity and the use of telescopes to learn about certain aspects of space). Despite these two marked differences, all texts will be comparable in terms of accessibility and general human interest, thus providing a solid ground for investigating the particular benefits of SDNs for understanding and remembering scientific information.

Chapter 3: Getting it right in the text: Developing high quality, accessible and engaging science texts for middle school students

Chapter Summary

In the second chapter, I described the science discovery narrative (SDN) and my process for understanding the underlying components of this type of scientific narrative. In this third chapter, I incorporate the SDN model into two of four experimental science texts that I developed for the purpose of determining the unique benefits of the SDN on understanding, recall and engagement for middle school students. The four experimental texts describe two different topics, chemistry and astronomy, and within each subject is a SDN and non-SDN text. Both text types (SDN and non-SDN) are especially structured to isolate the unique elements of SDN; thus all texts were created to have a consistent level of accessibility, coherence and interest for the targeted population of readers.

The development of these texts involved a 2-year process grounded in established theories about accessibility, coherence and engagement and driven by an iterative cycle of editing from scientists, literary experts, teachers and, most importantly, middle school students.

Introduction

The unique qualities of the SDN, as discussed in the previous chapter, provide the reader with a window through which the process of discovery can be viewed from the scientist's perspective. Several of the SDNs within my sample were intended for a general educated audience, as a part of the ever-growing volume of popular science literature. However, none of the SDNs was intended for a young reader, which is the target population for my study. In this chapter, I describe my process for developing four accessible, cohesive and engaging experimental science texts for middle school readers. Two of these experimental texts are SDNs related to two different science topics (chemistry and astronomy) while the other two texts are non-SDN versions of the same conceptual content. The non-SDN versions are similar to most school science textbooks in that there is no mention about the process that led to the conceptual facts that are presented. Only a brief mention of the scientist responsible for the information (e.g., *radium, which was first discovered by Marie Curie, is an element that . . .*) is included in the non-SDN texts. Generally, the information in non-SDN texts is presented as if the conceptual knowledge was always known, which is why I call these texts "conceptually known expositions" (CKEs). All four experimental texts are presented in Appendix A.

CKEs are similar to the content in most science textbooks in that both highlight the conceptual phenomena while ignoring the scientific process, but this is where the comparison ends. In order to determine the unique benefits of reading SDNs, it was important to develop a "control" text (CKE) that is comparable in every other way except for the SDN structure. Since I needed to control for the difference of the SDN, school science texts would not be an appropriate model for developing CKEs. As reported briefly in the second chapter, a text analysis conducted by the American Association for the Advancement of Science (AAAS, 2002) revealed a general sub-standard quality in the presentation of major scientific ideas. Classroom teachers, curriculum specialists and professors in science education were asked to rate the top nine best-selling science middle school textbooks in the U.S. according to multiple categories, which included the level of coverage of targeted concepts as well as the relevance and quality of conceptual descriptions. All nine textbooks consistently received a "poor" rating across the topic areas of earth science, life science and physical science. The raters noted a general lack of clear purpose and a failure to introduce terms and ideas in a meaningful, cohesive way; the raters also criticized the texts for attempting to cover far too many topics at the expense of depth into any particular concept. The issues raised in analysis by AAAS clearly go beyond a discussion about the need for including the process of discovery in school science texts. Thus, the CKE versions of my four experimental texts are not intended to emulate science textbooks but instead serve as comparative texts that convey the same concepts as the SDN in a clear and cohesive manner, but without the back story of the journey that led the scientists to their discovery. Thus, by controlling for conceptual clarity, they allow us to evaluate the "unique" qualities of the SDN.

The development of the experimental texts involved three general phases: a) selection of two different SDNs based on their alignment with targeted national and state content standards, b) development of age-appropriate SDN and CKE texts for the two selected science topics and c) an iterative process involving think-alouds with middle school students and subsequent edits of the drafted texts, which ultimately led to the final drafts to be used in the experimental study. There was a great deal of overlap in the development of all four experimental texts, but there was also a general systematic process in terms of addressing the qualities of conceptually accurate, accessible, cohesive and engaging texts, all of which I discuss in detail in this chapter. I also describe the editorial support I received from scientists and literary experts as well as the particularly valuable feedback from middle school teachers and students.

I made a conscious decision not to use any readability or text-analytic software during the actual development of these experimental texts; I believe that too often school texts are developed with the use of such software at the sacrifice of conceptual accuracy, coherence and meaningfulness to students. If one of my texts were to receive a readability score that placed it on the 10^{th} grade level, the natural impulse might involve one of two responses, a) break up the text into shorter sentences or b) replace low-frequency words (i.e., words that tend to appear less frequently in texts) with higher frequency synonyms. However, it is not clear that such editing practices will result in greater accessibility, even if these edits would change the readability level of the text to match the target reading level. Thus, any changes that were made to the four texts stemmed directly from the feedback I received from my target readers, scientific experts and literary editors. After the development of all four experimental texts, I used the online software program, *Coh-Metrix version 1.4* (McNamara, D.S., Louwerse, M.M., Cai, Z., & Graesser, A., 2005) to analyze readability, text cohesion and language according to a number of indices. Thus all the texts were ultimately "indexed" by readability metrics, but they were not "driven" by those metrics. This analysis will be discussed towards the end of this chapter.

Selecting Two SDNs

I selected Curie's (1904) *Radioactive Substances* and Galileo's (1610) *Starry Messenger* to serve as the model SDNs from which I developed conceptually and linguistically appropriate texts for the general population of middle school readers who are generally between the ages of 11 to 13 years of age (grade levels 6-8). My selection was first based on the national (National Committee on Science Education Standards and Assessment and National Research Council, 1996) and California state (California State Board of Education, 1998) science content standards because it was important that the content presented in the experimental texts aligned with the expected conceptual goals for middle school students, thus being a potentially useful activity for the participants in my study. Had I selected Guth's (1997) *Inflationary Universe*, for example, the conceptual information may have been too abstract for many students to understand without proper guidance, especially since theoretical physics is not addressed before high school. Additionally, had I selected Finaly y Barrés' (1881) account on yellow fever, the conceptual information may have been easy enough to comprehend, but there would be less utility in terms of addressing the content standards, which was an important benefit for participating teachers.

Curie's (1904) SDN provided a nice extension of the required content of the structural foundation of elements and very basic principles of chemical reactions. Similarly, Galileo's (1610) SDN complements the required standards of understanding the features of planets and stars as well as the underlying mechanisms of telescopes. Both accounts by Galileo and Curie (1904) were conceptually relevant for middle school science programs and thus were the winning candidates for the development of my experimental texts. The titles for the CKE and SDN versions of both topics are *Seeing at a Distance* and *Radioactivity*, respectively.

Composing Initial Drafts of SDNs and CKEs for Middle School Readers

Much goes into composing worthy science texts for middle school students. First, a writer must be mindful of conceptual and linguistic parsimony. Readers should not be overwhelmed by the amount of conceptual information presented in a science text, nor should there be too many unfamiliar words or phrases that would inhibit understanding, especially if those words or phrases are not providing critical information. However, it is equally important that concepts are presented in a way that fosters accurate understandings in order to avoid potential misconceptions that may inadvertently develop from the use of everyday language to describe concepts. Thus, there is a tension between accuracy and familiarity for middle school readers, and it is important that school science texts provide an optimal balance between these two qualities.

In addition to conceptual and linguistic parsimony, it is important that school science texts engage the reader in a meaningful way. First, the text must have some motivating qualities that encourage the reader to continue to turn the pages. Secondly, the text should present a purpose for learning the targeted concepts so that students learn about the practical benefits of acquiring such knowledge (AAAS, 2002). In other words, all of the experimental texts should contain an element of engagement and human interest, regardless of whether or not the process of discovery is explicated. The following points within this section address each of these qualities of my experimental texts.

The Appropriate Level of Conceptual Complexity

As previously mentioned, the national and California state science content standards served as a guide for targeting the particular concepts that would be highlighted in the experimental texts. My consulting scientists also had a great influence on the development of the conceptual framework that served as the foundational structure for the experimental texts. My lead consultant for the *Radioactivity* texts was Adam Mazur, a chemist for a pharmaceutical company in Cincinnati and co-discoverer of an amino acid that may eventually lead to the prevention of muscular atrophy. The lead consultant for the *Seeing at a Distance* texts was Dave Doggett, optics engineer in San Jose and inventor of numerous types of lenses. Both Mazur and Doggett were excellent sources of knowledge in that each had not only an exemplary understanding of his respective conceptual domain, but also a particularly extensive understanding for the respective scientists who had made the discoveries highlighted in the SDNs.

Within the boundaries set by the standards, I began to outline the corpus of conceptual principles with the guidance of my consulting scientists. Mazur, for example, provided detailed descriptions about the fundamental properties of radioactivity, identifying the core concepts that are the foundation of Curie's (1904) work. Similarly, Doggett provided a thorough explanation of concepts like "field of view" and the refractive properties of convex and cave lenses that Galileo (1610) created for his powerful "spy glass." After a couple of months of researching and identifying specific core concepts for both topics, I began to compose initial SDN and CKE drafts.

Generally speaking, the development of the SDNs was comparatively easier than the development of the CKEs in that the narrative structure provided a natural set of guidelines for composing the initial SDN drafts. I address the relative difficulty of composing the CKEs during the discussion about coherence below.

A Balance Between Brevity and Clarity

While working on the initial drafts, the tension between issues of brevity and clarity became increasingly more apparent. I continued to consult with Mazur and Dogett throughout this process in order to maintain conceptual accuracy while avoiding more explanation than was necessary. I planned to have students read two of the four experimental texts (an SDN of one topic, a CKE of another) within a single sitting (approximately 1 hour) in addition to answering a series of recall items about each text. Thus, I limited the length of the SDNs and CKEs according to the expected fluency levels for the middle school level (Hasbrouck & Tindal, 1991), which resulted in the guideline of limiting each text to about 1300 words or less.

With this general word limit in mind, Mazur and Doggett advised me (within their respective areas of expertise) which concepts should be discussed at length as well as which concepts to briefly mention in order to maximize the level of clarity presented in each of the four texts. For example, it was important to include a thoughtful review of what an element is and how they are the building blocks of matter in the *Radioactivity* texts in order to help readers understand the unique properties of radioactive elements. However, a detailed account of how the electrometer was built was less important than the general explanation of how the electrometer indirectly measured levels of radioactivity. Likewise, a complete account of Galileo's (1610) discovery of how the same side of the moon is always facing us would require a great deal of explanation about rate of the revolutions of the moon as it circles around the earth, the danger being that such an explanation might overshadow his major discovery of the moons of Jupiter and his argument that not everything revolves around the earth. As I continued my correspondence with my two consulting scientists, a general outline of key concepts surfaced for both topics.

As I continued to develop the first drafts of my experimental texts, another question related to the word limit surfaced—should both the SDN and CKE texts for each topic have the same word limit? Or is it more important to be consistent with the amount of conceptual descriptions, thus presenting a balanced emphasis on particular ideas regardless of the SDN structure? If the SDN version of *Radioactivity*, for example, presents a lengthy description of how Curie (1904) discovered the various qualities of radioactive elements, then this description should be fully accounted for, except for the actual process for finding such qualities (e.g., burning, rendering, etc.). If the radioactive element, polonium, is just briefly mentioned in the CKE version of this same topic, then it should receive the same brief mention in the SDN version. However, if the SDN and the CKE versions of each topic had the same word length, then the amount of conceptual description would be compromised in order to accommodate the additional wording required for the discovery process for the SDN version. Thus, consistency in the amount of actual conceptual information provided for each version of each topic was deemed more important than maintaining a standard word limit, making the SDN versions for both topics lengthier than the CKE counterparts. Some may argue that the greater length of the SDN versions provides the readers with greater time on task and thus more learning about the topic itself. However, this issue of "time on task" is, in a sense, the crux: whether or not the longer SDNs have a facilitative advantage for conceptual understanding, or if the additional story elements are distractions from the core conceptual information. Alternatively, one might argue that the additional reading burden of the SDNs might well tax the stamina of middle school readers, thus reducing their capacity to build cohesive models of meaning for the texts.

Conceptual mapping. In order to ensure the consistency in the amount of conceptual information presented across the two versions of a topic, I created a conceptual map for each of the texts according to the basic propositions (single-unit ideas) presented in each text. The process of conceptual mapping (i.e., outlining the network of single-unit ideas within a text) allowed me to identify portions of the text that needed greater or lesser detail in order to maintain consistency across versions. I then recruited the support of two researchers in the field of science education who were familiar with conceptual mapping, Lauren Barth-Cohen and Beat Schwendimann, to independently create conceptual maps for *Seeing at a Distance* and *Radioactivity*, respectively. These independent maps offered an objective eye for determining consistency across versions for each topic.

Accessibility and Coherence

It was important that all four of my experimental texts be grammatically and lexically accessible and cohesive because, as I explained earlier, the sole basis of comparison should be the SDN structure. To assist me with this task, I recruited the collaborative support of retired English professor from Riverside Community College, Gail Piestrup, one of whose specialties is children's literature. In this section, I describe this co-editing process, initially addressing the qualities of accessibility and coherence in turn.

Accessibility. In the development of student science texts, there is a tension between conceptual explicitness (which often requires more complex syntactic realizations and rare, concept-oriented vocabulary words) and linguistic simplicity (which generally requires less complex syntactic realizations and simpler vocabulary). A "simple" view of accessibility is most evident in readability formulas such as the *Flesch Ease of Reading* (Flesch, 1948), which holds that the fewer words in a sentence and the more familiar these words are (i.e., the more frequently particular words generally appear in texts), the less difficult it is for readers to comprehend this text (Klare, 1984; Zeno, Ivens, Millard, & Duvvuri, 1995). Word frequency is a good predictor of whether students will be able to read or offer a definition for a word, which is a crucial aspect of reading comprehension (NICHD, 2000; RAND Reading Study Group, 2002). Simply put, the more frequently a word occurs in a language, the greater the likelihood that students will know its meaning. However, high frequency words tend to denote more general concepts or categories (e,g., man or work rather than radiologist of employment); thus it may be argued that the more frequent a word, the greater the likelihood that, while students will know its meaning, this meaning may be less precise than what was intended in the text (Carey, 1985; Gopnik, 1996). This may especially be so with words in science, where a less frequent word such as "electrometer" conveys a level of precision that "measuring tool" does not.

Any account of text difficulty that uses sentence length to establish the readability of texts assumes, at least implicitly, that unpacking the ideas within a single, complex sentence is more difficult for readers than making connections across related propositions stated in separate sentences. A short sentence in itself may be easier to comprehend than a complex one. However, the challenge may come when the reader needs to integrate a cohesive meaning from a series of short sentences, which leads to greater demands on readers to make accurate inferences (Bowey, 1986; Pearson & Camperell, 1981). It is really a question of whether the memory burden of complex sentences trumps the inference demands of integrating ideas across separate propositions.

With regards to accessibility, Piestrup was an excellent sounding board; she guided me to an optimal level of explicitness (i.e., including key conceptual words, connective cues and other embedded clauses) for each of the four experimental texts. Piestrup encouraged me to focus on the goal of presenting the participants in my study with texts that not only teach something about science, but also model the kind of academic and literary language that we hope middle school students will eventually acquire. This goal for creating "model" science texts will be addressed further in relation to issues of coherence.

Cohesion. Cohesion, in this context, essentially means the "connectedness" of a text. If the words and sentences within a given text relate strongly to one another, then the text is considered as having a high level of coherence (Halliday and Hasan, 1976). For the SDNs, I had the added benefit of the discovery narrative structure that naturally allowed for causal and intentional constructions that contributed to the overall cohesion of a text (Miller, Beckwith, Fellbaum, Gross, & Miller, 1990; Van den Broek & Trabasso, 1996). With the guided support of Piestrup, I composed the SDNs with a particular focus on the natural flow of the unique components (i.e., love for the work, acknowledgement of the known, identification of the unknown and the experimental/explorative process toward discovery). Being mindful of the wording and the use of embedded clauses (as previously mentioned), the process of writing the SDNs for *Radioactivity* and *Seeing at a Distance* was, relatively speaking, a simple task.

Composing the CKEs proved to be a far more challenging task in that the explanations about the highlighted concepts needed to be presented in a logical order that maintained the "connectedness" of the ideas presented in these texts. The use of transitional phrases was especially crucial for helping the reader understand how one idea or concept related to another. Piestrup and I struggled along with my consulting scientists to determine the best logical order for presenting the concepts within the CKEs that maintained the overall coherence of the texts. For the CKE version of *Radioactivity*, for example, it seemed most logical to begin, after a brief introduction, with a general description of radium including a brief mention of Marie Curie, followed by an explanation of where it is found (pitchblende), its unique qualities, etc. This differed from the SDN version of radioactivity, which, after a brief introduction, began with an account of how one of Curie's colleagues, Henry Becquerel, discovered pitchblende, then followed with the series of events that led to the ultimate discovery of radium towards the end of the text. In both SDN and CKE texts, the amount of emphasis and explanation for each of the key concepts are similar, as I explained earlier. Thus, the overall goal was to present participating readers with science texts that are accessible and have a logical flow, regardless of version.

In order to have an external check on the overall accessibility and logical flow of the four experimental texts, I recruited the evaluative skills of experienced librarian Susie Goodin, who specializes in adolescent literacy. Generally speaking, Goodin believed that both pairs of texts (by topic) were balanced in their presentation of concepts and general quality writing, and agreed with the decision to omit the description about the rotational behavior of the moon, which was explained earlier. However, Goodin also commented on the complexity of terms within the versions for *Radioactivity*, suggesting that most students will not have been exposed to words like "radium" and "electrometer." Because these terms were repeated throughout the two texts, Goodin suggested that I leave the texts intact in order to see whether or not problems arise during the think alouds with middle school students (see section 3.4).

Multiple exposures to key conceptual words. Children need multiple exposures to less-frequently used vocabulary (i.e., academic vocabulary) in order to acquire word meanings form text (August, Carlo, Dressler & Snow, 2004; Nation, 1990 &2001; Stahl, 2003). Thus, in order to ensure the greatest possible level of accessibility, all four experimental texts contained multiple instances of key conceptual words (e.g., *radioactivity, refraction, concave,* etc.). Multiple uses of the same words will also improve the overall cohesion of the texts because the conceptual meaning will be further unified through the consistent presence of key vocabulary.

Narrative elements in CKE versions. While building a logical flow of ideas within the CKE texts, a question regarding the particular genre of these texts began to take shape. For the CKE texts, transitional structures were crucial for signaling a new idea about the given topic, and some of these structural features could be viewed as narrative-like. For example, the CKE version of *Radioactivity* contains the following transitional piece: "Radioactivity has the power to destroy lives. But in very small amounts and with full protection, radioactive elements can actually save lives in different ways." This description of the conflicting qualities of radioactive elements can be viewed as a "double-edged trope," which is a narrative construction that signals the evaluative aspect of a given text. Does this mean that the CKE texts represent a *different* kind of scientific narrative? My intentions have never been to determine differences in recall and interest for a narrative versus a non-narrative text. As stated earlier, my goal is to determine the unique benefits of one particular type of narrative text called the SDN. Based on my goal, the logical comparison would be a SDN versus a non-SDN, which is a particular type of text that is similar to the SDN text in every possible way save the unique SDN structure. Thus, although there may be some narrative element(s) in the non-SDN (CKE) text, this element is not a unique or defining feature of the SDN narrative.

Engagement and Human Interest

Composing conceptually appropriate, accessible cohesive science texts is important, but it is equally important to provide a reader (especially a young reader who may already have a disinterest in reading school science texts) with a plausible reason or sufficient motivation for reading such texts.

Engagement. Are narrative texts inherently more engaging than other types of texts? As I described in the first chapter, many researchers have argued against the use of narratives in science education, or at least fictionalized and/or "pseudo" narratives, in pressing for more non-fiction, informational texts that might be ultimately more engaging and incite further reading (Caswell & Duke, 1998; Worthy, Moorman, & Turner, 1999). However, it is clear that these researchers have not considered the narrative structure of many non-fiction, informational texts as illustrated in the first two chapters. Although readers may ultimately prefer the SDN version, it was important to me that the CKEs had some element of engagement that was similar to the SDNs in order to avoid a "straw man" competition between text types (i.e., exciting narrative about discovery compared to a boring exposition of facts).

Most of the research on reading and engagement relates to qualities of the reader or the instructional practices rather than qualities of the text itself. However, I found some research (Eccles, Wigfield & Schiefele, 1998; Guthrie & Wigfield, 2000; Wigfield, Eccles & Pintrich, 1996) helpful in thinking about the ways that both the SDN and CKE texts can encite an active engagement by directly addressing the readers in the texts. These ideas led to the final decision to have identical beginnings for both the SDN and CKE texts of the same topic; thus both begin with a few questions for the reader to consider, followed by a brief introduction to background information about the respective topic. The following introductory paragraph for the texts *Radioactivity* illustrate this point:

> What are the basic elements that make up the earth? If you could break down a rock into its smallest parts, what would you find? Gold? Copper? Elements are individual pieces of matter that combine with each other and make up everything we see around us. Most of what we see and use in the world, such as air and water, is not made of one single element. For example, sodium and chlorine are two different elements that make up the salt that we use for cooking.

For both the SDN and CKE, the reader begins with a series of questions to ponder. The use of pronouns "we" and "us" further pulls the reader into the discussion about elements.

Another aspect of engagement included formatting the SDN and CKE experimental texts in the form of individual, whole books with illustrative pictures. Rather than presenting the texts as a collection of 8.5" x 11" pages stapled together, I formatted each individual text to represent booklets that students could thumb through as they would a book. The pictures were identical for both the SDN and CKE texts for each topic, although they were placed differently according to their relevance to a particular part of the text (see Appendix A). My goal in formatting all of the experimental texts in this manner was to engage the participating readers in a task that mimicked the authentic practice of reading whole texts, thus focusing their attention on making meaning from the texts as one would naturally do in a library or bookstore rather than inadvertently framing the entire task as a typical "worksheet" classroom activity. Additional attempts to motivate the participating readers during the administration of the experimental study focused on the experimental tasks rather than the texts themselves (see Chapter 4).

Human Interest. The unique elements of SDN must be separate from what makes stories so interesting, so compelling. A narrative is inherently relevant because it is a primary means humans use to organize experience (Bruner, 1991), and thus it has the ability to pull the reader into the told events. SDNs allow the reader to experience the unfolding events along with the protagonist (in this case, the scientist). The struggles and excitement in discovering something new are captivating, especially if this discovery has an impact on human beings and the way they live. As I mentioned above, it is not compelling to compare a text that is an exciting account of discovery with a text that is a banal description of the facts that are the result of this discovery. The presence of human interest is not a defining feature of SDN. The account of a scientist's journey to discovery is a powerful structure for helping students understand the evolution of an idea, which in turn allows for deep conceptual processing and ultimate retention of the targeted phenomenon. In order to ensure that human interest does not vary significantly across versions, it was important to guarantee its presence within the CKE texts.

What is human interest? According to renowned philosopher Habermas (1972), science is strongly connected with the *technical* domain of human interest, which describes the human relevance of new knowledge about the world. The *technical* characterizes the perceived relevancy of new theoretical understandings as these theories develop into practical and valuable shared knowledge. With this theoretical perspective in mind, I included two qualities of human interest across the four experimental texts. The first of these qualities addresses the relative importance of a given topic and targeted concepts, essentially answering the often unspoken yet ever-present question of young readers, "why should I care about this subject?" Within the introductory paragraphs of each text, I addressed this question head on, engaging the readers with initial questions (as mentioned earlier) and providing a reason for continuing to read the text about the targeted concepts. Many young readers are naturally curious about our world and should thus easily relate to the general interest in radioactive elements as well as celestial phenomena.

The second quality of human interest addresses the practical importance of the targeted information in the experimental texts. For *Radioactivity*,

The Iterative process of Vetting and Editing Towards Final SDNs and CKEs

Once I developed the initial drafts of the SDNs and CKEs for both topics, I recruited 15 consented middle school students (grades 6-8) who live in Seattle, Washington to read my experimental texts and tell me what they thought in terms of what they understood, liked and disliked about the texts. I treated this process as a qualitative exploration in that my goal was to understand the thinking processes of my participating readers as they read a randomly assigned pair of texts that included a SDN and CKE from the different topics. Essentially, I asked the students to "think aloud" as they read, letting me know what they understood from the text or if they found anything confusing. All of the reading participants came from a single urban middle school in Seattle whose population is predominantly African-American. I selected this site for conducting think alouds because a) the students reflected the target population of urban middle school readers and b) as a former colleague to several of the teachers in this school, I was able to gain the initial access needed to obtain full consent from the school principal and participating readers.

The process of selecting participating readers was the full responsibility of the teachers who identified appropriate candidate readers after I explained the general parameters of my study and my desire to "try out" my experimental texts with students who reflect the general range of reading ability and scientific understanding of urban middle school students. The following table profiles the range of academic and language ability as well as ethnicity among the participants.

Table 3.1. Descriptive profile of participating readers for initial think-alouds

Total number of readers:		15
Grade Level:		
	6th grade	5
	7th grade	5
	8th grade	5
Gender:		
	Female	8
	Male	7
Academic status:		
	General Education	11
	Special Education	3
	Spectrum (Gifted & Talented Program)	1
Ethnicity:		
	African-American	6
	White	4
	Asian	4
	Other (Bi-racial)	1

I requested an equal number of 6th, 7th and 8th grade students in my think-aloud sample in order to get some sense of the kind of explanations and opinions students might offer across the three grade levels. The majority of the participating readers are considered to be general education students, which means that they are neither considered as advanced in their academic abilities nor as having any form of a learning disability. However, I also requested that a few of my participating readers be either in an advanced program or identified as having a learning disability so that I had the opportunity to investigate the full range of understanding that readers may have about the experimental texts.

Think Alouds

Think-aloud protocol. I interviewed all participating readers on an individual basis. After an initial informal chat about my project as well as their interests and thoughts about science and reading, I asked participants to read one of the randomly chosen texts in a manner that was best for them, which generally meant silently. After reading each page, I asked the participants to explain what they understood and if there was anything confusing about what they read. All but two students read two texts of different topics, one SDN and one CKE version. Once the students finished reading both texts, I asked which text they preferred and why.

I randomly selected two students to read all four experimental texts to determine any overall preference for the SDN version. I explained to these two readers that I wrote about two topics in two different ways, but I did not explain in what ways the two versions differed. My hope was to learn about how the readers would characterize the differences between the SDNs and CKEs.

I conducted all the interviews and made the conscious effort to present the project in an objective manner that would not lead the reader to consider any differences between the experimental texts. I explained to each participant that my project was about improving school science texts and that I wanted to know what he/she thought about two of the four science texts I have developed.

Response processes. For the most part, each of my readers had little difficulty understanding the selected texts. Within the first three readers, all minor typographical errors (10 total errors across all four texts) were identified and immediately corrected. Students usually read a paragraph or two at a time, either stopping to explain what they understood or to ask whether or not they could continue to the next page.

Most of the participant readers tended to read very "close" to the text in that they took time to read each word and often revisited earlier parts of the text or particular illustrations before proceeding to the next page. This behavior of careful reading did not seem surprising since a) I was present during the entire session and b) I had pointedly asked each reader to do his/her best in making sense of what they were reading and offering honest opinions about the quality and accessibility of these texts. However, the methodical reading behavior I observed during these interviews punctuated the importance of allowing an appropriate amount of time for reading during the actual study. On average, the think aloud interviews lasted twenty minutes; those identified as having learning disabilities needed more time, extending the interview session to thirty minutes. These initial findings on timing helped me gauge what would be required in the main study in terms of time. I knew that the reading portion of the main study would not include interrupted moments of discussion; students would therefore have as much time as they needed in order to read an entire text before having to respond to any questions. Thus, I concluded that the reading task for the study would take approximately half the time used for the interview, which would provide ample time for students to read and respond to subsequent measures during the main experimental study.

Thoughts and confusions about the "Radioactivity" texts. Readers generally had particular interest in the "life or death" qualities of radioactive elements, specifically, in the effects of radioactive poisoning discussed in both versions of this topic, their gaze lingering on the illustrations that showed hands swollen and marked from over exposure. It became clear that this part of both texts needed to be adjusted so that readers are not distracted from the conceptual information. Thus, I shortened the parts that described radioactivity poisoning and removed the picture with the swollen hands from both text versions, reasoning that it would serve as a "seductive detail" that would detract from the main purpose and for reading the selections.

Another revision resulted from the general confusion that the rock called "pitchblende" is also called "uranium ore." The two labels for the same substance seemed not only redundant but also distracting from the more important information about the composition of pitchblende. After a brief consultation with my science consultant (Mazur), I replaced all mention of "uranium ore" with "pitchblende" as it seemed to offer a clearer distinction between the radioactive rock of several elements (pitchblende) and the particular elements contained in this rock (uranium).

Thoughts and confusions about the "Seeing at a Distance" texts. For the most part, readers did not seem to have any issues with the wording of these texts. All readers seemed to be more familiar with telescopes and some aspects of planetary rotation than with concepts about radioactivity. Revisions in several of the illustrations were made in order to a) further clarify the direction of light as it travels through a lens towards the eye and b) provide more detailed anatomy of the eye and its parts. There are almost twice as many illustrations for the *Seeing at a Distance* texts than with the *Radioactivity* texts, which is most likely due to the nature of the topic, which is about viewing objects. Since both versions for each set of texts contain the same exact illustrations, the overall difference in the number of illustrations for each topic is inconsequential.

Preferences in texts. Nine of the thirteen participants who read only two texts (i.e., one SDN and one CKE of different topics) preferred the SDN version. However, reasons for preferring the SDN did not include any specific mention of the discovery story. Reasons for preferring this type of text included a general interest in the given subject matter (4 readers), an opportunity to learn more about a given subject (3 readers), and greater clarity in the explanation of the subject matter (2 readers).

Both students (one 7th grade male, one 8th grade female) who read all four texts preferred the SDNs over the CKEs. The 7th grader immediately noticed that the SDNs contain the story elements of the discovery and explained that the CKEs were "kind of boring, just facts and information" without a reason for learning this information. The 8th grade reader explained that she especially enjoyed learning about what happened to Curie, as if the text was a mystery novel.

Conclusions: The Final Version of the Experimental Texts

High Quality Across All Four Versions

I focused on a theoretically grounded, qualitative and co-constructed approach for developing four experimental texts worthy of being read by middle school students. My full confidence in the quality of the scientific content and linguistic presentation of all four texts, regardless of genre, was crucial for two reasons. First, it is important that all the middle school students who participated in my study were exposed to high quality texts from which they could learn about particular science concepts. Second, it is important that my study is a comparison of genre (i.e., SDN vs. CKE) and that all other qualities are held constant, including the quality of writing. I am confident that all appropriate measures were taken to ensure quality across all four texts.

Once the four texts were in final draft form, I recruited four middle school teachers to evaluate all four texts based on their general quality of content and engagement. Two of these teachers were located in Seattle and taught 6th grade language arts and 7th grade special education, respectively, at the time of the study. The other two teachers resided in the San Fransisco Bay area and taught 7th and 8th grade science, respectively. Only one of the four teachers (8th grade science) believed that the SDN versions would elicit greater levels of engagement and thus result in greater levels of understanding and memory recall of conceptual information. The remaining three teachers believed that all four texts had an equal chance of engaging students, depending on individual interests in particular science content. One teacher (7th grade science) stated that those students with a particular interest in science might prefer the CKE versions in that the story elements of the SDN texts would be more distracting and thus less appealing.

Coh-Metrix Analysis

After the completion of the experimental texts drafts, I submitted the final versions to be analyzed by the free online *Coh-Metrix* software program (McNamara, D.S., Louwerse, M.M., Cai, Z., & Graesser, A., 2005). In this program, the texts are analyzed according to six general characteristics, which include general reference information, readability indices, general word and text organization indices, syntactic complexity, referential and semantic indices and "situation model" dimensions.

Much of the data produced from the Coh-Metrix indices reveals a fair amount of similarity for both versions of texts within each topic (see Appendix B for the full report). However, it is unclear what would constitute a significant difference as descriptions of each index provide a general continuum of "more-to-less" of a particular quality with an occasional commentary on the favorability of such qualities. Further, the differences in text length make type-token ratios less meaningful for determining a relative ease of reading. Acknowledging these limitations in using Coh-Metrix data for characterizing the experimental texts, the following indices offer interesting evidence about the texts. These data suggest a high degree of linguistic similarity between the SDN and CKE versions of each topical pair.

Table 3.2. Coh-Metrix Indices

Title	READFKGL 'Flesch-Kincaid Grade Level (0-12)'	INTEi 'Incidence of intentional actions, events, and particles.'	LSAassa 'LSA, Sentence to Sentence, adjacent, mean'	LSApssa 'LSA, sentences, all combinations, mean'	LSAppa 'LSA, Paragraph to Paragraph, mean'
Radioactivity (CKE)	10.081	8.351	0.484	0.39	0.653
Radioactivity (SDN)	10.103	10.717	0.427	0.379	0.649
Seeing at a Distance (CKE)	7.879	18.388	0.525	0.521	0.652
Seeing at a Distance (SDN)	7.952	18.769	0.513	0.508	0.636

Readability. According to the *Flesch-Kincaid* readability scale, the *Radioactivity* texts should be much more difficult for my readers than the versions for *Seeing at a Distance.* This difference in readability is not surprising considering the general low frequency of words like *radioactivity*, *electrometer* and *radium* compared with words *lens*, *convex* and *rotation*. The important point here, in my opinion, is that the two versions for each topic have about the same level of readability and thus serve as worthy comparisons for the study.

Incidence of intentional actions, events and particles. According to the Coh-Metrix tool, presence of "intent" is marked by a main verb that is performed by an animate subject, the protagonist, and is associated with forms of narrative texts. Based on this description alone, I had predicted that the SDNs would have a significantly greater incidence of intentional constructions since the discovery narrative is focused on a protagonist (the scientist) doing things such as asking questions, performing observations or experiments, etc. Although both SDNs have a higher value in this dimension, the difference in the presence of intentional constructions within the *Seeing at a Distance* texts seems rather slight compared to the *Radioactivity* texts. The CKE version of *Seeing at a Distance* seems to have a greater focus on the reader (who essentially plays the role of protagonist, although not as a discoverer, but as a student who is exploring established knowledge) than the CKE version of *Radioactivity*. Thus, it is possible that the relative lack of intentional constructions for the *Radioactivity* CKE will have a negative impact on the conceptual recall and interest for this text.

Latent Semantic Analyses (LSA). LSA relates to the "conceptual similarity" of the words, sentences and paragraphs within a given text. The closer two words are related to one another (e.g., *cat* and *dog* are more closely associated with one another than *cat* and *cow*), the more semantically cohesive they become. In all three categories within the LSA dimension, the SDNs have a consistently lower value compared with the CKE versions. Again, it is unclear whether these differences are significant enough to warrant attention, but the consistency in lower values suggests that the narrative texts are less semantically cohesive, providing some supportive evidence for the idea that narrative texts are distracting rather than facilitative for learning conceptual information.

Next Steps

In this chapter, I described in detail the process for developing four high quality science texts that differed in topic (radioactivity and planetary phenomena) and type (SDN and CKE). In the next chapter, I describe a randomized controlled study examining differences in relevant outcomes for the four experimental texts.

Chapter 4: Effects of the Science Discovery Narrative (SDN) on the Understanding of and Interest for Middle School Science Texts

Chapter Summary

In the third chapter, I described the complex process used to develop four high-quality middle school science texts. In this fourth chapter, I report the purpose, methods and findings of a randomized, controlled experimental study in which I compare how well two different types (SDN and CKE) of science texts elicited interest, conceptual understanding and long term recall from 7th and 8th grade students. During this study, participants (N=209) were randomly assigned to read two of four science texts that differed according to topic (*Radioactivity* and *Seeing at a Distance*) and type of text (SDN and CKE), and then answered a series of questions about the conceptual information presented in each text. A week later, participants answered another series of questions about these same two topics in order to determine the longer-term effects of a particular type of science text.

Findings from this comparative analysis showed that participants recalled a greater level of conceptual information when they read the SDN version of a given text. This result was present for each recall task except for the first recall task for *Radioactivity*. Additional analysis showed an interaction between ethnicity and treatment, indicating that the effect of receiving the SDN was significantly greater for minority students for all except the final recall task for *Seeing at a Distance*. Also, for 8th grade participants, the difference between the SDN and CKE versions of *Radioactivity* was non-significant, which might well be explained by the reported high level of prior conceptual knowledge among 8th grade students on this topic.

These findings in favor of the SDN suggest that the discovery narrative is important for those with less prior knowledge about a given concept in that it helps in building a contextual relevancy that promotes understanding and long-term memory of key conceptual information. Further, there is moderate evidence to suggest that students of marginalized ethnicities respond even more positively to the SDN, thus further punctuating the importance of including the discovery process when learning about science through texts.

Introduction to the Problem

Research in the field of science literacy has kept a steady spotlight on issues of genre in school science texts. Many have argued, for example, that including narrative texts (i.e., stories related to science in some way) in the science curriculum not only encourages the inadvertent development of misconceptions, but also limits the opportunity for further exposure to expository texts -- namely, science textbooks that are more descriptive and less story-like in nature (Duke, 2000; Jetton, 1994; Madrazo, 1997; Mayer, 1995; Rice, 2002; Guzzetti, Williams, Skeels, and Wu, 1997). However, others have argued that narrative texts have a greater power to reveal the true nature of science as a process, promoting engagement and conceptual understanding (Baram-Tsabari & Yarden, 2005; Landau, 1991; Lemke, 1990; Snow, 1963; Rosenblum & Markovits, 1976; Sohmer & Michaels, 2004). This present study attempts to a) clarify some of the fog that has settled on this debate on narrative and science and b) determine if there is any significant educational benefit for the use of narrative in science education.

From a broader, more theoretical perspective, it seems that the controversy about using narrative texts in the science curriculum at least partially stems from differing views on the definition of "narrative." As I elaborated in the first chapter, narrative encompasses many types of experiential accounts, and some of these types naturally exist within the field of science (Falk & Yarden, 2009; Landau, 1991; Latour & Strum, 1986). Thus, for this study, I focus on one specific type of narrative, which I call the "scientific discovery narrative" (SDN). After investigating numerous first-hand written accounts of discovery (SDNs) by scientists (see Chapter 2 for a detailed account of this investigation), I adapted two of these SDN's for the middle school audience. Specifically, I adapted Marie Curie's (1904) *Radioactive Substances* and Galileo's (1610) *Starry Messenger*, two first-hand SDNs, to reflect the general reading level and conceptual relevancy of middle school science curricula and created versions of each that are somewhat comparable to what is often read in school science textbooks, a type that I call "conceptually known expositions" or CKEs (see Chapter 3 for a detailed account of the process in developing these SDN and CKE experimental texts). With these texts, I hope to answer the following questions: a) what effect does the Scientific Discovery Narrative (SDN) have on a young reader's ability to recall conceptual information? b) What effect does this particular genre of narrative have on a reader's interest in a given science topic?

The reason for focusing on the SDN is two-fold. First, as is stated earlier, there is some confusion about the precise definition of "narrative." Stories that highlight the authentic process of science (e.g., a scientist's exploration under the ocean) are either unintentionally grouped with fantasy narratives that may only allude to scientific information (e.g., a talking snail that is scared of deforestation) or are regarded as nonfiction and thus grouped with textbooks and other "factual" sources. Choosing to focus on the SDN in this study will hopefully broaden the current debate and ultimately foster a deeper discussion about genre in science literacy education.

The second reason for focusing on the SDN is the overwhelming need for science educational texts to address the evolution of scientific processes underlying conceptual information. This neglect of scientific process in most contemporary science texts may partially account for the fact that students leave school with little more than decontextualized databases of knowledge accumulated through fact-memorization rather than deep conceptual understanding (Bowker, 2005; Niaz & Rodriguez, 2000), let alone any appreciation of the dispositions that characterize the scientific process and the personal attributes that scientists bring to their work. Many cognitive scientists have echoed this argument by pointing out a lack of understanding about the evolutionary process of scientific knowledge (i.e., the nature of science as a way of knowing) among elementary and secondary students and teachers alike (Ackerson & Abd-El-Khalick, 2005; Gallagher, 1991; Griffiths & Barman, 1995; Lederman, 1992; Moss, 2001). Thus, my goal in this study is to investigate the comparative advantages of the SDN, which may serve as a response to the neglect of scientific process in science education.

Methods

Participants

Participating schools. 209 7th and 8th grade students participated in this study.[1] Participants and their parents signed the approved forms of consent, a process that was administered by the classroom teachers. These students come from nine classrooms in two different middle schools located in northern California. The two schools were recruited through an online, widely distributed advertisement. Initially, 248 students submitted personal and parental consent, but only 209 participants participated in the study due to absenteeism, which is discussed later.

Four of the nine classrooms come from a school that I will refer to as "Oaks Middle School," which is considered a high-poverty urban school with a predominately Latino (60%) and African American (34%) population. Approximately 23% of Oaks students are performing at or above proficiency level in science, and only 17% are considered proficient or above in English-language arts.

The remaining five classrooms come from a suburban school that I will refer to as "Ocean Middle School," which, by contrast, is a low-poverty school with a predominately White (64%) and Latino (30%) population. Approximately 72% of Ocean students are performing at or above proficiency level in science, and about 62% are proficient or above in English-language arts.

Student participants. As I reported earlier, 209 out of 248 students with consent participated in this study. Prior to the administration of the study, I collected background information from each participant, which included demographics (gender, grade, ethnicity, language proficiency and home language), overall reading performance (i.e., state English-language arts standardized score provided by the schools, and teacher-ranked overall performance in science class. This last variable was provided by each of the participating teachers, who ranked students within each respective class according to level of ability in science. I used this background information to a) confirm a consistent balance in ability and demographic make up across all treatments, b) investigate interactions between the demographic information and performance and c) evaluate the associations between the experimental measures and the external measures (i.e., state ELA score and science ranking performance score).

The ethnic background of these participants reflects the overall profile of the respective schools. However, the average academic performance (i.e., proficiency level for English-language arts) of the participants is relatively higher than the respective school average. Individual, standardized science proficiency scores were unavailable for this study. The state-level, middle school science exam is only given at the 8th grade level, which occurred a few weeks prior to this study.

---

[1] All research methods were approved under IRB protocol 2008-3-4, by the Office for the Protection of Human Subjects (CPHS) at the University of California, Berkeley. Initial approval was granted on May 6, 2008, with continued approval until May 6, 2010.

Table 4.1. Demographic information of participants (*N*=209), by school

	Oaks Middle Participants	Ocean Middle Participants
Ethnicity:		
African American	18 (28%)	2 (1%)
Latino	42 (66%)	33 (23%)
Other	4 (6%)	6 (6%)
White	0 (0%)	106 (73%)
Proficient or above in English-language arts	43 (61%)	128 (78%)
Total	64 (31%)	145 (69%)

Prior knowledge of participants. According to the 7th grade teachers from both participating schools, their students have either very little or no recent exposure to the major concepts presented in experimental texts. The 7th grade teacher from Oaks Middle expressed considerable doubt about whether her students had any background knowledge about the nature of elements or the principles of refraction even though they are expected benchmarks for the state and national content standards (California State Board of Education, 1998; National Committee on Science Education Standards and Assessment and National Research Council, 1996). However, the 8th grade teacher explained that a considerable portion of the science curriculum for her students centers on the periodic table, chemical reactions and the structural aspects of elements. This teacher also explained that she recently had given a short presentation about Marie Curie's discovery of radium. Thus, it is entirely possible that the 8th grade participants in my study already possessed enough knowledge of the targeted concepts in both versions of *Radioactivity* that there will be no significant effect of version on recall performance.

Missing data. Due largely to absenteeism, only 209 of the total 248 participants with consent were present during the first session of the study, which involved reading the experimental texts for this study. Those participants that were absent during this first session but present during the second session were excluded from participating, since they would be unable to answer any follow-up questions about the experimental texts.

I conducted both sessions of this study during the final month of the school year (June 2009), during a time in which both participating schools experienced an increase in absences. Participating classrooms from Oaks Middle had experienced the most severe attrition in student attendance, with 30 absences compared with only 9 absences from Ocean Middle. Of the 209 participants who were present during the first phase (i.e., the session in which all students read the experimental texts and then answered a series of questions about the conceptual information of these texts), 192 participants were present during the second phase (the sessions in which students answered another series of conceptually based questions about the targeted topics). A total of 17 participants who were present during the first phase were absent during the second phase, 10 from Ocean Middle and 7 from Oaks Middle.

Materials

The texts. Four texts were created for this study—two versions (SDN and CKE) for two topics (*Radioactivity* and *Seeing at a Distance*). The process of creating these four texts spanned a two-year time period, which involved extensive research into primary science texts and multiple iterations of drafting, editing (by scientists, linguists and literary experts) and think-aloud sessions (with middle school children of varying reading abilities). The composition of all four texts was founded on national and California state standards (National Committee on Science Education Standards and Assessment and National Research Council, 1996; California State Board of Education, 1998), theoretical frameworks regarding coherence, genre and human interest (Habermas, 1972; Halliday and Hasan, 1976; Miller, Beckwith, Fellbaum, Gross, & Miller, 1990; Van den Broek & Trabasso, 1996) and empirical research on readability and text complexity (Bowey, 1986; Carey, 1985; Gopnik, 1996; Flesch, 1948; Hasbrouck & Tindal, 1991, Klare, 1984; Pearson & Camperell, 1981, Zeno, Ivens, Millard, & Duvvuri, 1995). All four texts were formatted into little books in order to provide the participating readers a context resembling the actual practice of reading whole texts.

I drafted the two SDN texts according to the four major components of first-hand written accounts of discovery by scientists (i.e., love for the work, the known, the unknown and the experimental or observational process). Initial SDN drafts were first edited by scientists for accuracy and then further edited by a children's literature professor for coherence, engagement and accessibility. Conceptual maps of both pairs of texts were created in order to clarify the presence and organization of the conceptual and narrative propositional unit-based information in these texts. The goal in developing these texts was to create two text versions for each of the two topics that were similar in every possible way (in terms of accessibility, cohesion and human interest) except for the unique qualities of the SDN structure. Based on feedback from all participating editors and students mentioned above, the final versions met this goal. A more detailed description of this process of text development is provided in the third chapter.

The measures. All the measures created for this study focus on either a) understanding and subsequent recall performance of conceptual information presented in the experimental texts or b) interest in the conceptual information presented in the experimental texts. I describe the process of developing the conceptual recall and interest tasks in turn.

Understanding and recall. All recall items were deliberately designed to focus on the conceptual information and excluded any explicit narrative elements (e.g., Marie Curie contracting radioactive poisoning). I applied the methodology of the *4 building blocks approach* (Wilson, 2005), which begins with building a construct map of the varying levels of the targeted constructs in the texts. In order to best represent the essential task of reading science texts, my construct map includes varying levels of text comprehension loosely based on Kintsch's (1998) *Construction-Integration (C-I)* model of reading comprehension.

Kintsch's C-I model. According to Kintsch, an individual approaches a text not as a blank slate, but as a reader with some prior understanding about a given topic that is highlighted in the text. Thus, when we read, we integrate new knowledge from the text with our prior knowledge, which builds a revised understanding about the given topic. Kintsch calls this revised understanding the "situation model," which may represent more than the text, depending on the reader's level of prior knowledge. The less a reader knows about the topic described in a text, the more "closely" she needs to read this text. For new topics, a reader may need to read very carefully, even rereading certain parts in order to build an accurate understanding of the text. The more a reader knows about the information presented in a text, the less she relies on the text itself, even skimming over certain sections as a confirmation of prior knowledge.

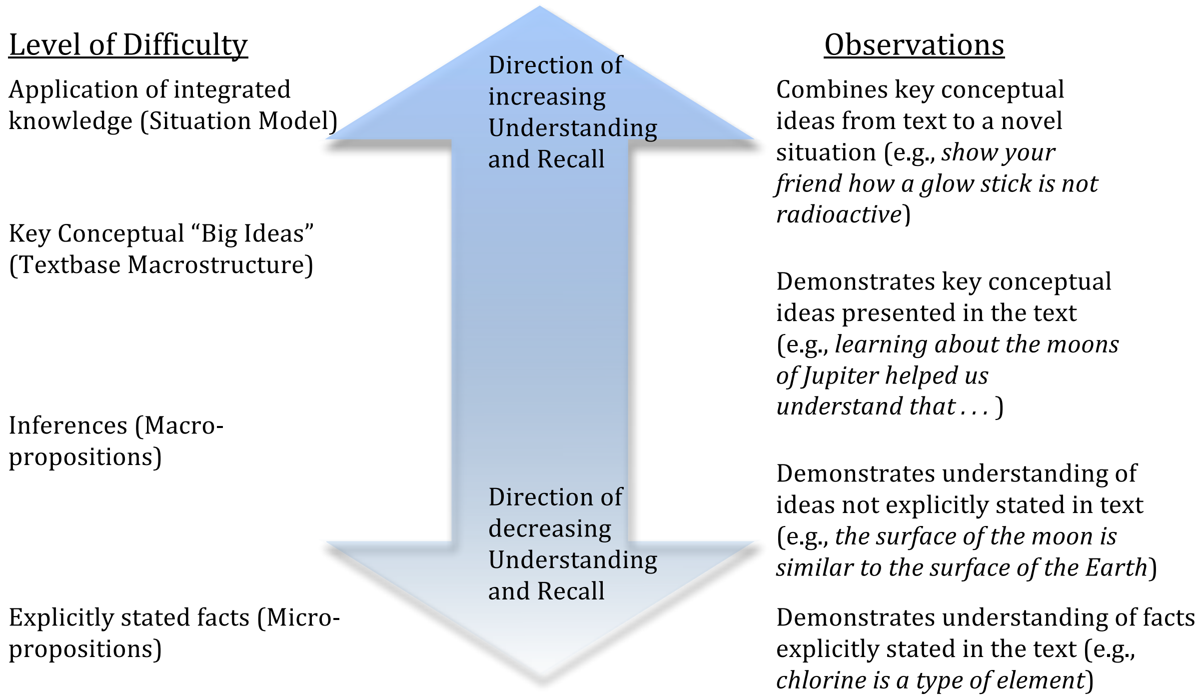
Kintsch explains that every text is made up of distinct ideas that he calls "micro-propositions." Consider the following example: *Harry gave Sam a red rose*. There are at least three distinct ideas within this single sentence. First, Harry gave Sam something. Second, this "something" was a rose. Third, this rose was red. These are *concrete* ideas in that they are explicitly stated in the text. But there is more presented in this sentence than these explicit micro-propositions. We can make several inferences from Harry's gesture, the first being that Harry at least feels enough regard for Sam to give a rose to this person. Another inference is a potentially amiable connection between these two people. When we connect the micro-propositions in this sentence with micro-propositions of previously read sentences, we build bigger ideas about Harry and Sam, which inform the reader about the significance of the rose. Kintsch calls these bigger ideas "macro-propositions," which comprise the key points of a given text.

The micro and macro-structural components of a text create the "textbase." As the reader builds a text-based understanding about Harry and Sam, she not only learns more about their relationship and the significance of the rose, but also she is able to form her own ideas about this relationship, which is the situation model. Kintsch describes this model of reading comprehension as a cohesive process that does not purposely imply differentiating levels of cognitive power. In other words, there is no sense from Kintsch's work that building a textbase of macro-propositions is more cognitively difficult than understanding the distinct micro-propositions presented explicitly in the text. Nor does Kintsch suggest that the pinnacle of cognitive difficulty is the situation model, which is the integration of the textbase and prior knowledge. My use of Kintsch's C-I model goes beyond its intended use in order to theoretically ground the varying levels of information presented in my experimental texts and tapped by the comprehension and recall tasks given to the participants.

Construct mapping the experimental texts. According to Wilson (2005), it is important to be theoretically deliberate in developing measures in order to provide evidence for validity based on the test content. What are the different levels of understanding of these experimental texts that a reader can demonstrate? An effective recall measure for each of the topics in my study depends on its ability to capture the range of understanding and recall of conceptual information that would be expected from a diverse sample of middle school students. Thus, each of my measures must contain items that vary in difficulty in order to address this expected range in ability. Wilson explains that these levels of difficulty should be deliberately defined (or rather mapped out) and founded on an established theoretical framework. The defined levels of a given topic comprise what Wilson calls a "construct map."

A construct map is designed to explicitly describe the varying levels of a given construct and also describe the expected demonstrations of the expected ability at each level. I have developed the construct maps for the two science topics presented in this present study (*Radioactivity* and *Seeing at a Distance*) with the consideration of Kintsch's (1998) C-I model because I believe this model serves as the best possible framework for building a measure for a science text. My goal is not to understand the general levels of understanding a given concept, like radioactivity; the experimental treatments are directly tied to the texts and not necessarily the presented topics. Although all four texts are well developed in that they have been carefully crafted and edited by several experts, they still only represent a portion of the constructs and do not fully present the entire conceptual framework of the scientific information. The C-I model, then, serves as a super-structural framework for the levels of understanding and subsequent recall of the conceptual information presented in the experimental texts. The following figure presents the construct map used for developing items for both topics of texts.

Figure 4.1. Construct map of understanding and recall of experimental texts



Construct maps typically describe the level of difficulty (above left) and the kind of observed response associated with each level (above right) along a continuum of the construct. Admittedly, this is an over-simplified view of reading comprehension in that there is much to consider about the act of reading a text that go well beyond a single continuum of difficulty. However, this simplification is a necessary step in deliberately creating an array of items that capture at least some of the varying levels of understanding and recall that readers of these texts are expected to demonstrate. This construct map is based on the notion that the more information one has to integrate in order to successfully respond to an item, the more difficult that item will be. Thus, items that require the application of major, macro-structural ideas to novel situations as described for the highest level (situation model) are considered most difficult.

Based on this construct map, I developed multiple-choice and short answer items (see Appendix C) that attempt to address each of the levels of difficulty. Items that highlight facts explicitly stated in the text were logically deemed easier than such items that required more inference making. Further, distinct, micro-structural ideas were considered easier to comprehend and identify than macro-structural key conceptual ideas. Based on this assumption, the focusing power of a concave and convex lens within a telescope (presented in both versions of *Seeing at a Distance*) is more difficult to comprehend than the general structural properties of a single type of lens.

It is important to acknowledge that not every macro-structural idea is necessarily more difficult than describing facts (which would more likely fall into the microstructure) explicitly stated in the text. A title of a book is a reader's first cue to the kinds of information that might be contained inside the pages, and often incites the reader to consider the general ideas that may be presented. These generalized orientations are helpful to the reader, activating a schematic representation that facilitates reading (Bransford, 1985; Norris & Phillips, 1987). However, these general schematic notions (e.g., *Seeing at a distance* describing what it is like out in space) are not the targeted ideas for the construct map presented above. I will revisit this issue when I discuss the validity of my recall measures within the data analysis section below.

It is equally important to note that the initial items of all recall tasks (and the final item for recall tasks for the first session) were structured primarily to gauge the interests of the readers and not necessarily designed to assess a level of recall described in the construct map above. These interest-related items will be discussed in further detail in the following section.

Editing process of measures. All measures went through the same rigorous process of editing as the experimental texts. Scientists, literary experts and teachers evaluated and gave feedback on the content, format and wording of each item. I performed think aloud interviews with the 15 student readers described in the third chapter. Each reader read, responded to and commented on the relative ease or difficulty of each item, which helped to confirm the validity of the construct map. Based on feedback from the think aloud interviews, I made final edits to the measures. Items and tasks related to interest went through this same editing process.

Interest*.* Embedded in the recall tasks were four different items intended to gauge the participant's level of interest toward the assigned experimental texts. During the first session, participants were asked the following three questions:

- (For both topics, first item in Recall I task): *What did you find most interesting about this text?*
- (For both topics, second item in Recall I task): *Do you have any questions of your own about this topic? Write these questions below.*
- (Last item after Recall I task): *Which book was most interesting to you? (circle one): telescopes radioactivity Explain your choice.*

The second session included the following question:

- (For both topics, first item in Recall II task): *Have you tried to learn more about radioactivity since you last read this text? (circle one): YES NO*

  Write what you have done to learn more.

These interest-related items served two purposes. First, they helped to establish a non-evaluative context for the recall task in that my purpose for the recall items is not to "test" the students but rather evaluate the effectiveness of the text as a source of conceptual information. Second, these interest-related items are helpful in determining the level of engagement and interest of readers for the different versions of science text, thus potentially providing further evidence of any differences in performance across types of texts for a single topic.

Interest Ranking Task. In addition to the embedded interest items presented above, participants completed an interest-ranking task to determine general interest in various topics in relation to science-related topics (see Appendix C). This ranking task was given twice at the very beginning of both sessions. Participants were asked to select 10 topics from the list provided that they would most like to read about and rank them according to their preference. Thus, a participant must make two decisions about each of the topics listed. First, she must decide whether or not a particular topic should be included within her top 10 and second, she must decide the position (if selected) of this topic. The purpose of this ranking task is to compare the relative interest of particular science topics (e.g., *radium*) with other more popularized topics (*fashion*) and determine whether this relative interest changed after reading the experimental texts.

The list of topics was presented in non-numbered rows, and each participant received these topics in random order to avoid the potential confound of order. A broad collection of specific and general topics were included on this list, a small number of which were related to the text topics (i.e., *Jupiter, radium, science, telescopes*).

Traditionally, researchers across many different fields of study have used Likert-scale items for measuring attitudes. However, in light of reported issues with comparing attitude across different groups and lack of clarity in the definition of the construct (e.g., the confound between extremity of a position and the intensity with which it is held) I decided against the use of a Likert-scale task for measuring attitudes (Clason & Dormody, 1994; Converse & Presser, 1986; Goldstein & Hersen, 1984). Although the ranking task provides only a qualitative look at patterns in and general shifts across selected topics, I believe that the information will be more useful and concise for comparing responses across different treatment groups and across time (i.e., changes in selected topics from the first to the second session).

Study Procedures

I administered most of the procedures for this study. All data collection occurred during the participants' science class period, and the respective science teacher was present the entire time of data collection. All tasks were simultaneously administered to the participants. As I mentioned earlier, only those participants with consent that were present during the first session continued to the second recall session, which occurred a week later.

Prior to the first session, I visited each participating classroom to introduce my project. I was mindful of how I framed this study in order to avoid what has been termed "stereotype threat." According to social psychologists Steele and Aronson (1995), one of the consequences of stereotype threat is the effect of academic pressure on test performance by minority students. The performance of minority students tends to suffer when they perceive that their academic performance is being evaluated. Thus, my introduction to all the participants in my study was the same to all classes:

> "I want you to take a look at these books, read them, and then answer some questions about them afterward so that I can learn about what you thought about them. Your answers will help me know if the books were any help in learning something about science."

I avoided any technical terminology (*comprehension ability, evaluation*, etc.) that might inadvertently "raise the stakes" for the recall measures. I believe I was successful in that throughout the following sessions, many students felt comfortable enough to approach me to share additional thoughts about what they liked or disliked about the science texts as well as offering advice on many aspects that ranged from the given content to the use of visual graphics.

During this introductory session, I introduced the interest-ranking task described earlier, which reinforced my non-evaluative position as a researcher. Once all consent letters were collected, classroom teachers administered the ranking activity prior to the first session.

Session I: Administration of texts and first set of recall measures. Participants with consent who were present during the first session of the study were randomly assigned to read two experimental texts, each a different topic (*Radioactivity* and *Seeing at a Distance*) and genre (SDN or CKE). Participants read each administered text in turn, and each individual reading was followed by a series of questions about the conceptual information presented in the previous text. Participants were unable to look back at the texts while responding to the questions, the content of which was described earlier. The following example items come from the first recall task for *Radioactivity* (see Appendix C for first and second recall items for both topics).

Item 6: Radioactive elements are decaying, which means that:

a.) they are slowly dying out
b.) they are generally unchanged
c.) they are losing parts of atoms
d.) they are gaining parts of atoms

Item 3: Your friend shows you a glow stick and thinks it is radioactive because it glows. Explain at least **<u>two ways</u>** to show your friend that the glow stick is not radioactive. **(open-ended)**

Thus it is probably better to regard these items as measures of immediate knowledge retention rather than comprehension items (since most measures of comprehension allow students free and continuous access to the text).

In order to account for the order in which the topic and genre was read, the texts were administered according to the following conditions:

Table 4.2. Text assignment configuration for all four conditions

	***Radioactivity*: SDN**	***Seeing at a Distance*: CKE**	***Seeing at a Distance*: SDN**	***Radioactivity*: CKE**
1st Reading	**A**	**C**	**D**	**B**
2nd Reading	**C**	**A**	**B**	**D**

Each participant was randomly assigned to one of these four conditions presented in the table above (A, B, C, or D), which indicate the topic and type of texts that were read during the first and second readings. The order in which the SDN version was read for a given topic was counterbalanced in order to control for potential order effects.

The following table presents the distribution of ELA proficiency, gender and ethnicities across all treatments:

Table 4.3. Distribution of abilities, gender and ethnicity by treatment (N=209)

	Treatment A	Treatment B	Treatment C	Treatment D
Mean standardized score English-language arts (with standard deviation)	365 (58)	366 (59)	369 (50)	370 (60)
Gender (number):				
Females	26	24	32	25
Males	29	29	18	26
Ethnicity (number):				
African American	6	7	3	4
Latino	17	16	19	21
Other	4	0	1	2
White	28	29	25	24

Before participants read the texts, I administered the interest-ranking task described earlier, which asked each participant to choose from a random list of 20 topics their top 10 favorite reading topics. Each participant received a unique randomized list of the same 20 topics. Recall measures were administered directly after participants read each of the two texts. Subsequent interest ranking and recall tasks (also described below) were administered to each participant a week later.

Although I administered the texts and all response tasks during class time, I made great efforts to frame this study as an activity that positioned me as the learner and the participants as helpers who offer feedback. I was consistent with my introductions and explanations for each participating class, telling the students that I was interested in learning about what they thought about the texts because of my desire to help make improvements in the science texts that students like themselves read in school. I emphasized the point that the recall tasks were not "tests," but a way to see if the text was memorable enough, that if they couldn't remember or didn't know, it would be good information for me as a researcher. These instructions were reinforced during the second recall session.

Session II: Subsequent ranking task and recall measures. A week later, I administered another set of the interest-ranking tasks to each participant with consent who was present during the previous session. Following the interest-ranking task, I administered in random order the second set of knowledge retention items for both topics. Participants did not revisit the texts during this session.

As I stressed in the previous visits, I reminded the students of their role as evaluators and my need to understand how well the texts helped them remember important information in the texts.

Scoring the Measures

Along with the support of two additional researchers, I entered all of the data for each participant and coded the open-ended, short answer responses. Approximately 20% of the responses to the constructed response items were scored by all three raters, and an additional 30% were scored by two raters. After initial training sessions, within this sample, interrater reliability exceeded.90. Any disputes in scores were resolved through discussion.

Analyses of Scales

The fourth step in Wilson's (2005) building block framework (as mentioned earlier) is the measurement modeling phase, in which a relationship between the construct and the items is formalized through a statistical model. In this case, I have used the Rasch model (Rasch, 1960) for modeling my constructed measures. The Rasch model provides interpretive tools, such as the Wright map, which provides a quality control check on the degree to which the expectations of the instrument design process, as outlined earlier, have been met.

Reliability

In order to show the internal consistency of the experimental measures, estimated person-separation reliability coefficients (i.e., EAP reliability as estimated by the Rasch model) are given in the table below.

Table 4.4. Person-separation reliability coefficients for measures

	Number of Items	Person-Separation Reliability
Radioactivity, Total items	37	.86
Radioactivity, Recall I	22	.80
Radioactivity, Recall II	15	.69
Seeing at a Distance, Total items	17	.79
Seeing at a Distance, Recall I	8	.62
Seeing at a Distance, Recall II	9	.69

The greater apparent reliability for *Radioactivity* could be explained in part by the greater number of items within these measures. The bundle of items requiring participants to circle all elements (or radioactive elements) in the recall measures for *Radioactivity* (16 items total, see Appendix C) can partially account for this increase.

Relations with Other Variables

State standardized scores. The California state English-language arts test (CST-ELA; estimated reliability = .95; Educational Testing Service, 2008) is a general measure of academic ability and purports to measure vocabulary knowledge, fluency, comprehension, literary response and writing (California Department of Education, 1998). Similarly, the texts and measures for this study are academic in nature and require participants to understand key conceptual vocabulary, comprehend the major ideas presented in the texts and express some of these major ideas in writing. However, the CST-ELA test targets a different set of constructs than are measured in this study. Further, the "high-stakes" nature of the ELA is a great contrast to the more casual presentation of the experimental texts and measures, as I described earlier in the study procedures section. Thus, the moderate similarities and striking differences between the CST-ELA and experimental measures would seem to promise a moderate correlation.

After applying the formula for dissattenuation for measurement error, the correlation between the CST-ELA score and the *Radioactivity* total test (for both sessions) ability was estimated as r= .74, and the correlation between CST-ELA score and *Seeing at a Distance* total test ability was estimated as r= .62.

Science ranking scores. As mentioned earlier, all teachers were asked to rank order the students according to his/her respective classroom performance. Teachers ranked the students according to their demonstrated knowledge about the various topics covered during class. The correlations between the science ranking score and the total test ability for *Radioactivity* and *Seeing at a Distance* are r=.37

and r=.54, respectively. These correlations are not dissattenuated for measurement error, as it is not possible to estimate the reliability of the teacher rankings in the present study. One can infer that coefficients presented here underestimate true correlations by a small to moderate amount. The (raw) correlation between the CST-ELA and science ranking ability is r= .50.

Goodness of Item Fit

Once all data were entered and coded, I used the Rasch modeling software program *Construct Map* (Kennedy, Wilson & Draney, 2005) to model the coded recall items that measured understanding of the conceptual information presented in the text. All items that demonstrated a "good fit" to the Rasch model (i.e., infit meansquares between .75 and 1.33) were included in the data analysis. Below are the Wright Maps, which demonstrate this goodness of fit between the abilities of the participants (left) and difficulty of items (right), which are interpreted more fully below.

Figure 4.2. Wright Map for *Radioactivity* items

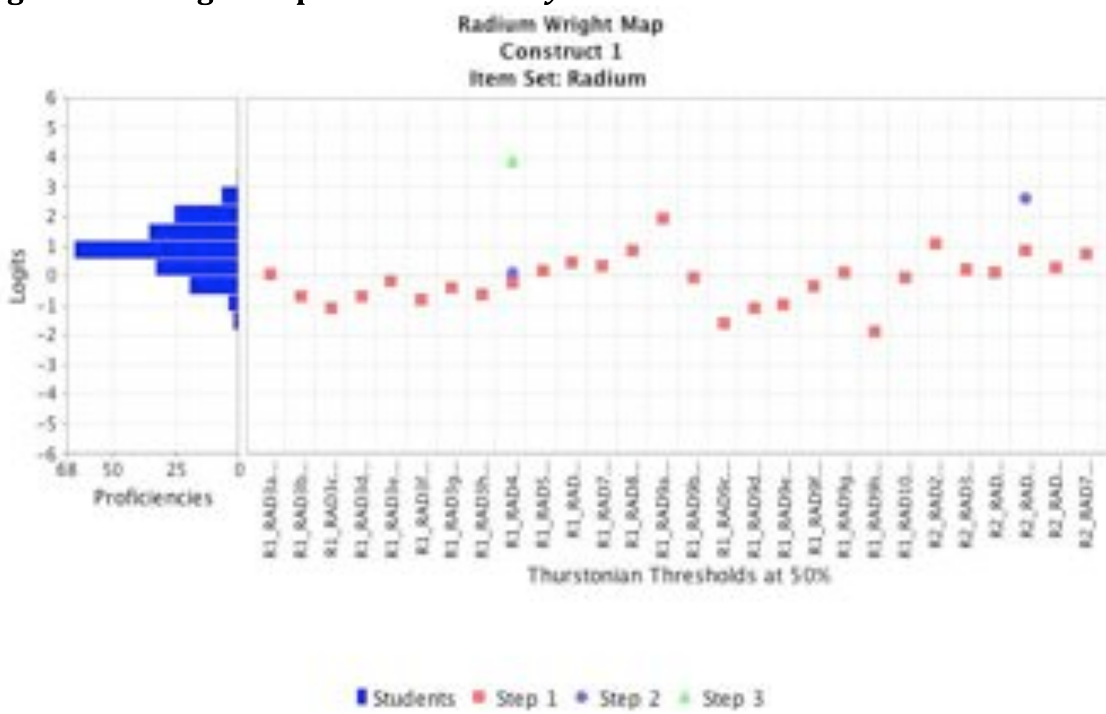
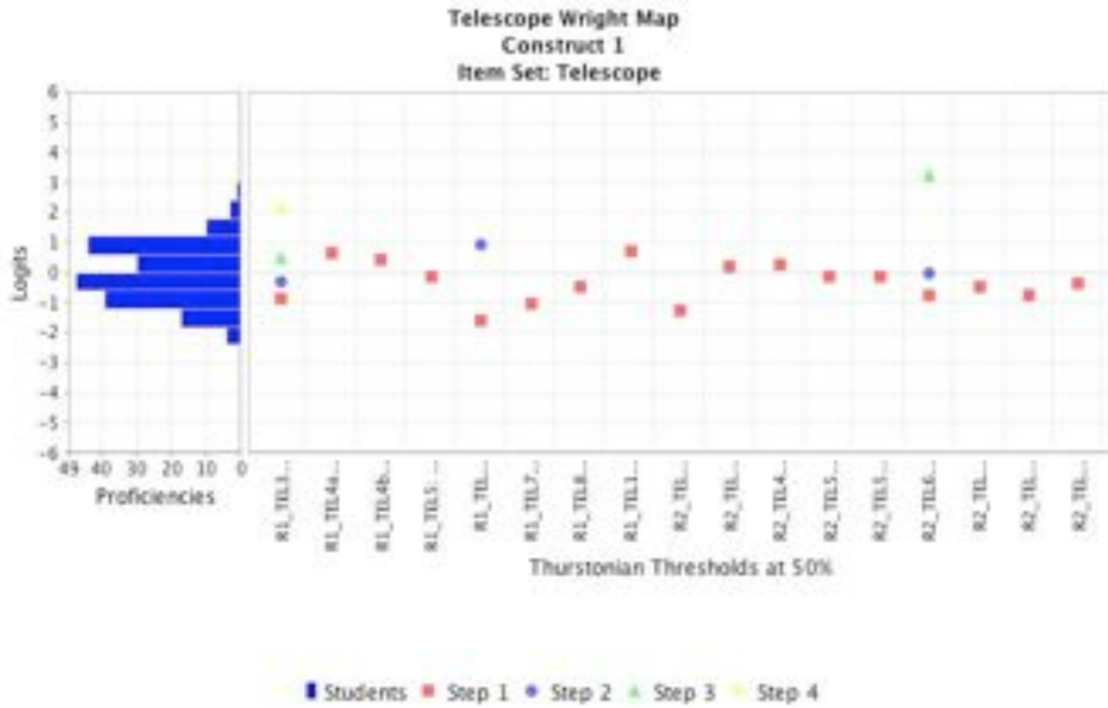


Figure 4.3. Wright Map for *Seeing at a Distance* items



Excluded item. During the coding period, it became abundantly clear that item number nine in the first recall measure for *Seeing at a Distance* failed to elicit a meaningful response from any of the participants (see Appendix C). This item asked the participant to choose which length of telescope (long or short) would be best for viewing objects out in space, followed by a request to explain their choice. Responses revealed a lack of understanding about the item and did not demonstrate any depth of understanding (e.g., *because I would be able to see farther*; *because the longer telescope is more powerful*). Thus, this item was discarded prior to the fit analysis since it failed to capture conceptual understanding about importance of tube length for telescopes (i.e., *a longer tube would allow for greater distance between the two lenses, thus resulting in greater magnification*).

Misfitting Items. The second question from both first-recall tasks did not fit (mean squares significantly above 1.33), which asked whether the reader had any questions about the text (see Appendix C). These were the only items that did not require a response from the participant, and further, they were less direct than other items in their measurement of the construct (i.e., an inference was made about the student's level of recall based on the quality of their question) than other items, which directly required students to demonstrate recall of text information.

Checking the expectations of the construct map. Examining the correspondence between the Wright Map and construct map helps to determine whether the intended difficulty of levels of the construct map that I discussed earlier effectively described the observed variation in item difficulty. As predicted, items that required the integration of key conceptual ideas for answering major conceptual ideas (especially in novel contexts) were most difficult for participants, such as item four in the first recall task for *Radioactivity* and item 6 in the second recall task for *Seeing at a Distance*. Item number eight in the first recall task for *Seeing at a Distance* required students to make an inference, which was generally more difficult than determining whether or not particular substances were examples of elements (bundled items for number three in the first recall task for *Radioactivity*), which was information explicitly provided in the text.

As I explained earlier, the creation and use of the construct map serves as tool for building a theoretically grounded array of items that will best capture the range of abilities of the participants in my study. The most important question is whether or not the responses from all participating readers reveal a unique benefit of the SDN versions of text on initial and long-term recall performance.

Results

In order to address my research question about the effect of SDN on recall performance of and interest for science conceptual information, I conducted two different kinds of analyses: a) a latent regression to examine the effect of genre on initial and subsequent recall performance for the two topics and b) a chi-square test of independence to examine the association between preference of text and treatment (i.e., whether the participants tend to prefer the topic of the SDN version rather than the CKE version). The initial and subsequent interest-ranking tasks served as supplementary data for a qualitative exploration into patterns of preference changes over time.

Recall Performance

For all recall data, I fit a series of latent regression models (Adams, Wilson & Wu, 1997) to the data, comparing those who read the SDN with those who read the CKE for each topic. Traditionally, regression assumes that the dependent variable is measured without error. In this case, the dependent variable is an estimate of a latent construct and is measured with error. Thus, I used a latent variable regression in the program, *ConQuest* (Wu, Adams, Wilson & Haldane, 2007). Instead of using the standard two-step procedure in which student scores are first estimated and then analyzed, the latent regressions used in this study directly estimate the effects of predictors on the latent variable without first producing individual scores. Thus, measurement error is built into the regression rather than ignored as in the two-step procedure.

The following tables present the estimated mean differences among the latent variables (on the original logit scale) for all recall tasks for the full sample.

Table 4.5. Estimated mean differences on recall, between treatment conditions (N=209)

Recall Performance (by Task)	Mean Difference in Person Location (SDN-CKE)
Recall I *Radioactivity*	.19 (.15)
Recall II *Radioactivity*	.35 (.13)*
Recall I *Seeing at a Distance*	.84 (.13)**
Recall II *Seeing at a Distance*	1.31 (.11)**

*p<.05; **p<.01. Standard errors are in parentheses.

Overall, there was an effect of SDN on performance for all recall tasks except the first recall task for *Radioactivity*.

Interactions. Due to potential differences in prior knowledge by grade as well as potential additional effects of treatment on various demographic variables, I investigated the interactions between the effect of treatment and all other demographic variables (i.e., school, grade, gender, ethnicity and language proficiency). This level of investigation explores the idea that exposure to the SDN could exert a stronger influence on recall performance for certain groups of participants.

Interactions between gender and language proficiency were not significant. However, there were significant interactions between treatment and grade, school and ethnicity. The following tables present the interaction between treatment and grade (Table 4.6), school (Table 4.7) and ethnicity (Table 4.8).

Interaction by grade. The following table presents the estimated mean difference in scale score for all four recall tasks, for 7th grade participants (first estimate column), followed by the estimated mean difference between 8th grade and 7th grade participants (second estimate column) and additional effect of treatment for 8th grade students (final column).

Table 4.6. Estimated main effect of treatment and estimated additional effect by grade

Recall Performance (by Task)	Treatment (SDN-CKE)	Grade Level (7th grade=0, 8th grade=1)	Grade by Treatment
Recall I *Radioactivity*	.21 (.18)	.71 (.20)**	.04 (.28)
Recall II *Radioactivity*	.64 (.15)**	.93 (.16)**	-.66 (.24)**
Recall I *Seeing at a Distance*	.79 (.16)**	.46 (.17)**	-.04 (.25)
Recall II *Seeing at a Distance*	1.35 (.15)**	.34 (.16)**	.13 (.23)

*p<.05; **p<.01. Standard errors are in parentheses.

Results show a negative grade by treatment interaction for the second recall of *Radioactivity*. This effect (-.66) is of approximately the same magnitude as the main effect of treatment (.64); thus, it appears there were no effects of treatment for either *Radioactivity* recall task for the 8th grade participants. The 7th grade participants, however, did benefit from the SDN version for this topic.

Interaction by school. The following table presents estimates analogous to the previous table, replacing grade level with school assignment. Again, the final column highlights estimates for "school by treatment" interactions.

Table 4.7. Estimated main effect of treatment and estimated additional effect by school

Recall Performance (by Task)	Treatment (SDN-CKE)	School (Ocean=0, Oaks=1)	School by Treatment
Recall I *Radioactivity*	.02 (.06)	-1.18 (.19)**	.55 (25)*
Recall II *Radioactivity*	.08 (.13)	-1.18 (.17)**	.91 (.25)**
Recall I *Seeing at a Distance*	.67 (.14)**	- .42 (.18)**	.52 (.26)*
Recall II *Seeing at a Distance*	1.20 (.13)**	- .36 (.17)**	.36 (.25)

*p<.05; **p<.01. Standard errors are in parentheses.

The first three recall tasks reveal significant interactions of school by treatment. Specifically, participants from Oaks Middle received greater benefit from the SDN version for both recall tasks for *Radioactivity* and the first recall task for *Seeing at a Distance* compared with participants from Ocean Middle.

Interaction by ethnicity. The following table presents estimates similar to the previous two tables, instead highlighting ethnicity. The final column highlights estimates for ethnicity by treatment interactions.

Table 4.8. Estimated main effect of treatment and estimated additional effect by ethnicity

Recall Performance (by Task)	Treatment (SDN-CKE)	Ethnicity (White=0, All other ethnicities=1)	Ethnicity by Treatment
Recall I *Radioactivity*	.06 (.06)	-1.19 (.17)**	.65 (.20)**
Recall II *Radioactivity*	- .15 (.15)	-1.20 (.15)**	.98 (.22)**
Recall I *Seeing at a Distance*	.91 (.17)**	- .53 (.17)**	.16 (.25)
Recall II *Seeing at a Distance*	1.51 (.16)**	- .62 (.16)**	.40 (.23)

*p<.05; **p<.01. Standard errors are in parentheses.

Participants who were members of underrepresented minority groups (i.e., mainly, in these schools, Latino and African-American) uniquely benefitted from the SDN version for the first and second knowledge recognition tasks of *Radioactivity*. The lack of sufficient numbers of participants within each ethnic category prevented further investigations into specific effects by ethnicity (i.e., looking at unique benefits for Latino participants compared with African-American participants).

The significant interactions presented here provide some evidence that students with little prior knowledge as well as students of low-economic status may benefit more greatly from the SDN version.

Interest

Results for participant responses concerning their interests involved three different kinds of investigations. First, I performed two separate chi-square analyses to a) determine the association between treatment and whether students indicated that they performed independent research on each of the two topics during the second recall session and b) determine the association between preference of text and treatment, as explained earlier. Second, I looked for patterns in the written explanations about what participants found most interesting about each text that they read. Finally, I explored the responses for the interest-ranking task to determine notable changes in the “Top 10” selection from the first to the second session.

Individual research. Only 32 participants reported that they did some individual research on any of the two topics, and there were no significant associations between these responses and treatment. Explanations about this individual research included internet searches, reading other texts about the topic and asking other individuals. One participant wrote that she consulted her physician about the x-ray machine (she was getting an x-ray of a recently broken foot) and her new knowledge about radioactivity. One interesting note is that based on the total 32 individuals who reported conducting some form of independent research, 21 of these responses were about the topic, *Radioactivity*. This finding suggests that the topic of radioactivity might have been more motivating for the participants, which will be addressed more fully during the discussion section.

Preference of text. The following table presents a cross tabulation of preference by treatment and the results of a chi-square test of independence.

Table 4.9. Chi-square statistics of preference by treatment

Preference	Those Who Read SDN Version of *Radioactivity*	Those Who Read SDN Version of *Seeing at a Distance*	Total
Radioactivity	70 (66.67%)	49 (47.12%)	119 (56.94%)
Seeing at a Distance	32 (30.48%)	49 (47.12%)	81 (38.76%)
No preference	3 (2.86%)	6 (5.77%)	9 (4.31%)
Total	105 (100%)	104 (100%)	209
Pearson $\chi^2_{(2)}$= 8.2692, p = 0.016			

Participants who read the SDN version of *Radioactivity* preferred this text over the CKE version of *Seeing at a Distance* almost two-thirds of the time. Those who read the SDN version of *Seeing at a Distance* preferred this text over the CKE version of *Radioactivity* about half the time. It is also important to note that there were no significant differences in preference according to gender, ethnicity or language proficiency. Thus, subject matter and text type were the only two variables that had an effect on text preference.

Narrative responses about points of interest. There were no distinct patterns in responses regarding particular points of interest for the texts. Responses ranged from various pieces of conceptual information (e.g., *it's cool how light bends to our eye*), applications (*I liked learning about the x-ray machine*) or the impact on people (*radioactivity can kill you, but it can also save your life*).

A notable point is the number of readers of the SDN of a particular topic who mention the discovery story as a point of interest in the text. Thirty-one of the total 105 readers of the SDN version of *Radioactivity* mentioned the Curies' story as a highlight and 48 of the total 104 readers of the SDN version of *Seeing at a Distance* found Galileo's story to be a point of interest.

Interest Ranking Task. Due to the enormous variability in the ranked preferences for the selected topics, I could not determine any significant discernable patterns. I was able to note that the number of times key topics related to the topic appeared on a participant's top 10 interest list increased between the first and second administration for participants who received the relevant SDN versions, such as "Jupiter" (six additional times as opposed to three fewer for readers of the CKE version of *Radioactivity*) and "radioactivity" (five additional times as opposed to fifteen fewer for readers of the CKE version of *Seeing at a Distance*). However, it is difficult to interpret these results definitively.

Discussion

Limitations of this Study

I composed the experimental texts and measures used for this study, and I was present for all the data collection, which certainly suggests the possibility of some form of bias in the results. However, I made every effort to guarantee a fair comparison of the SDN and CKE, the process of which was detailed in the previous chapter. I believe that all four experimental texts are worthy school science reading materials in that each was written accurately and cohesively with the intent of engaging middle school students.

The measures did not address the unique elements of the SDN versions, thus avoiding bias in performance. Although my presence throughout the data collection could not be avoided, I made every effort to focus my comments on all four texts and not single out any version or topic.

Another limitation of this study is the lack of an adequate external source of validity evidence for the experimental measures. Future replications of this study should include a more detailed and comprehensive evaluation of the validity argument for any outcome measures used. Nevertheless, I believe that the documentation of the test construction provided in this study and the quantitative evidence given above provide strong preliminary support for the proposed interpretations of the measures used.

Discussion of Findings

Results from this present study provide evidence in support of the hypothesis that middle school students benefit from texts that highlight the process of discovery, especially when prior knowledge is lacking. Results from the second recall tasks also suggest that the SDN supports retention of science conceptual information over time. Granted, the SDN did not always elicit greater knowledge retention (i.e., for the first recognition task on *Radioactivity* for all participants and the second recognition task on the same topic for 8^{th} graders) and it did interact with many characteristics of the participants (e.g., ethnicity and school attended), but the overall story here is that students learned and remembered more information from the SDNs.

The importance of the discovery narrative is further punctuated by the interactions for ethnicity and school assignment, which suggest that students from low socio-economic backgrounds benefit more from reading science texts that highlight discovery. Due to the fact that most of the participants identified as ethnic minorities (i.e., Latino or African-American) came from Oaks Middle, a reported high-poverty school, it is not possible to separate ethnic minority status and socio-economic status for this study. Future replications of this study could address this confound.

The overall pattern from the findings on recall tasks suggests a unique benefit for the SDN, but there is, as I suggested earlier, some inconsistency, especially for the first recall of *Radioactivity*. However, I believe that this inconsistency can in part be explained by the results of the various interest measures. Based on the chi-square analysis on text preferences, it seems that *Radioactivity* was overall more popular than *Seeing at a Distance*, regardless of version. Twenty-one percent fewer readers preferred the SDN text when the topic was about telescopes. Also, a greater portion (although this number is not significant enough to make a very strong claim) of those who read the SDN version of *Seeing at a Distance* had no preference between the two texts (see Table 4.9, page 74). This interest for *Radioactivity* is also evidenced in the greater number of participants who reported additional independent research on this topic. These findings suggest that the subject of radioactivity may be inherently more interesting, thus suggesting that participants paid a great deal of attention to this text, regardless of treatment. Because of this focused attention on *Radioactivity* texts, participants may not have received additional benefit from the SDN version.

The popularity of *Radioactivity* is particularly surprising considering that the level readability was much higher (approximately at the 10^{th} grade level) for the texts about radioactivity than for the texts about telescopes (7^{th} grade level, see the end of Chapter 3 for the structural analysis of the experimental texts). One could easily assume that a more difficult text (i.e., text with a greater number of unfamiliar words) would be less interesting, especially to young readers. However, it seems that the novelty of learning about a powerful and potentially dangerous element like radium is enough to capture the attention and interest of participants in this study, regardless of the readability level.

Conclusion

I began this chapter by presenting research in support of the argument for teaching the scientific process in that students and teachers alike have little knowledge about science as a way of learning something about our world. Findings from this present study contribute to this current research by providing evidence for the unique benefits of the discovery narrative on the retention and motivation for conceptual information presented in texts. Further, the evidence that suggests the unique benefits of the discovery narrative to be especially beneficial to students of low socio-economic or ethnic minority status.

This present study also serves as a call for clarification in the current discussion about academic texts in the field of literacy. Specifically, I mentioned the lack of clarity in the current debate regarding genre and school science texts. Dichotomizing texts into “narrative” and “non-narrative” categories neglects the established, theoretical narrative frameworks that encompass different kinds of narrative, discovery being only one of such kinds. Future investigations involving issues of genre should consider this broader perspective of narrative.

Chapter 5: Concluding Remarks

This dissertation study addresses two outstanding problems about science education. The first problem is the insufficient understanding about the nature of science as a process and an evolving body of knowledge among students and teachers in science classrooms. The second problem is the lack of support given by school science textbooks to rectify this first problem. In the first chapter, I suggested that both problems are integral to the current debate about the inclusion of narratives, or stories, in science texts by linking the process of science (i.e. exploration or experimentation) with the fundamental elements of narrative texts. I described how narrative is a human construct that people use, intentionally or not, to tell about many different types of experiences, including scientific inquiry. I provided a series of example texts that broaden the view of narrative that is traditionally discussed in education and explored the ways in which narrative is, and should be, an integral part of texts produced in science.

Since the earliest recorded times, scientists have written their accounts of how they came to know something new about the world. In the second chapter, I described my process for analyzing 12 discovery accounts in order to determine the commonalities across these texts. Science discovery narratives (SDNs) are similar to published scientific studies in that both describe a problem within a given field and the process that led to a solution, which is the discovery. However, SDNs are quite different from other scientific writing due largely to the scientist's prerogative to disclose the serendipitous events that occurred during the discovery process, which are typically either excluded from or reclaimed as standard procedure in published studies. SDNs, to a greater extent, show the scientist's love for the work as it relates to a given problem or mystery within a particular field of study.

SDNs vary greatly in detail and emphasis across four common components—the love for the work, the known, the unknown and the process towards discovery. While some include greater detail about the accounts involving mishaps and chance (e.g., Birkeland, 1913; Marshall, 2002), others emphasize the frustrations and jubilations they experienced during the exploratory or experimental process (Galileo, 1610; Guth, 1997). Although SDNs do not represent a consistent form of scientific writing, they offer a valuable lens for discussing many aspects of scientific work that are currently ignored in the classroom. Specifically, SDNs can help students understand the great variability in the culture of scientific learning as well as the human qualities of scientists themselves.

Through various accounts of discovery, students and teachers will have the opportunity to see that the scientific process is not a clean set of procedures, and that sometimes mistakes are even more valuable than planned experiments. Also, scientists can be completely wrong in their thinking about a problem (e.g., Finlay y Barrés, 1881) and need to start all over again. Essentially, SDNs emphasize the human aspects of scientific endeavors, which I believe provides a relevant, engaging context that further facilitates conceptual understanding. Thus, for this study, I investigated the SDN's unique effect on the interest, understanding and knowledge recognition of middle school students.

In the third chapter, I described how, in order to investigate the unique effects of the SDN, I undertook the systematic development of a series of experimental texts that are consistent in terms of accessibility, cohesion and general engagement (i.e., the presence of human interest) and differ only in terms of whether the unique components of the SDN were present in the text. These experimental texts were developed with the guidance of scientists, literary experts, teachers and students to differ in terms of genre (SDN and non-SDN versions) and topic (radioactivity and the use of telescopes to learn about certain aspects of space).

Identifying the common elements in SDNs was the first step in the process of adapting and developing the experimental texts for middle school students. These four common elements (love for the work, the known, the unknown and the process towards discovery) served as a guide for ensuring that the experimental SDNs are comparable to the first-hand accounts yet written at a level appropriate for young readers with little to no prior knowledge of the targeted scientific information. The non-SDN comparison texts are comparable to the prototypical SDNs in every way, save the four common elements. The scientific information presented in the non-SDN texts includes explanations of the targeted concepts without the discovery process, which inadvertently may give the impression that this scientific information was always known to the respective field of study, hence their name, conceptually known expositions (CKEs).

In order to maintain consistency in the quality of all four experimental texts, I used a theoretically grounded, collaborative approach for developing the SDN and CKE text versions for each topic. The rationale for this theoretically deliberate and complex writing process is two-fold. First, it is critical (and ethical) that the middle school participants in my study are exposed to high quality texts that can facilitate conceptual learning. Second, it is important that my study is only a comparison of genre (i.e., SDN vs. CKE) while holding all other qualities constant, which includes the quality of writing. Thus, multiple iterations of outlining, drafting and editing commenced with a steady focus on maintaining accessibility and cohesion throughout all four texts. This iterative process involved the editorial support of scientists as well as literary specialists to ensure conceptual accuracy and accountability for developing cohesive and engaging texts across genres.

A final evaluative process for all four experimental texts involved a series of think-alouds with 15 middle school participants who helped to further refine the texts in terms of general accessibility and cohesion. Additionally, I asked four middle school teachers to individually evaluate the four experimental texts, all of who agreed that all texts were sufficiently consistent in terms of quality and engagement. However, there was disagreement among these teachers in terms of which genre would elicit greater levels of engagement and thus result in greater levels of understanding and recall of conceptual information. Only one teacher believed that the SDN versions would elicit greater recognition performance. The remaining three teachers concluded that all four texts had an equal chance of engaging students, depending on individual interests in particular science content. However, one of the three in agreement postulated that those students with a particular interest in science might find the story elements of the SDN texts to be distracting and thus less appealing.

The difference in opinions of these teachers reaffirmed the consistency in the general quality of all four experimental texts. With the exception of the unique components of the SDN, all texts are comparable in terms of accessibility and general human interest, thus providing a solid ground for investigating the particular benefits of SDNs for understanding and remembering scientific information.

In the fourth chapter, I described the randomized, controlled experimental study in which I compared how well two different genres (SDN and CKE) of science texts elicited interest, conceptual understanding and long term knowledge recognition from participating middle school students. During this study, 209 7th and 8th grade participants were randomly assigned to read two of four science texts that differed according to topic (*Radioactivity* and *Seeing at a Distance*) and type of text (SDN and CKE), and then answered a series of questions about the conceptual information presented in each text. A week later, participants answered another series of questions about these same two topics in order to determine the effects of a particular type of science text on longer term knowledge recognition.

Participants generally recalled a significantly greater level of conceptual information for both the first and second tasks of knowledge recognition when they read the SDN version of *Seeing at a Distance*. This same result was present for the second recognition task for *Radioactivity*. The difference between the SDN and CKE versions of *Radioactivity* was non-significant for 8th grade participants, potentially due to the reported high level of prior conceptual knowledge among 8th grade students on this topic. Additionally, the effect of the SDN was significantly greater for participants of underrepresented minority groups (i.e., mainly, in these schools, Latino and African-American) on all recognition tasks except the second recognition task for *Seeing at a Distance*.

These findings in favor of the SDN suggest two points. First, the SDN is an important learning tool for those with less prior knowledge about a given concept. Presenting conceptual information within the context of discovery promotes understanding and long-term memory of key conceptual information. Second, the interactions for ethnicity suggest that students of marginalized ethnicities benefit even more from the SDN, thus strengthening the argument for including the discovery process in school science texts.

These findings prompt me to leave readers of this dissertation with two final thoughts, the first of which focuses on the take away points from the empirical study while the second expands this discussion to future research endeavors.

First, the essential point from the empirical study presented in Chapter 4 is that students benefit when they learn about how scientists learn about our world and have the opportunity to read texts that expose the historical progression of scientific knowledge. To only present the knowledge about what we think is true about the world is not truly science but merely a collection of information. However, I am not advocating for a complete revolution in the genre of school science texts. Not every text has to be an SDN. However, when students have the chance to read about scientific concepts in their historical and personal context, science comes to life, and that excitement may inspire extended readings of more technically descriptive information within the domain. Perhaps many of the students who reported engaging in further independent research after the first encounter with the SDNs were reading expository texts about the conceptual information presented in the experimental texts. The discovery narrative elevated both the importance and meaning of the highlighted concepts, which in turn made expository texts about these concepts important and meaningful to the students.

Second, our schools, teachers, and students desperately need better quality school science texts. It took approximately two years of research and development to create the texts for this study. Granted, much of this time was devoted to understanding the theoretical underpinnings of the SDN, and then operationalizing the theory through many drafts of the experimental texts, but unless our profession is willing to devote extensive time to the development and empirical evaluation of science texts, our students will continue to read and regurgitate the fact-laden, interest-barren texts that dominate the current marketplace of science textbooks.

As a recent entrant into the educational research community, I believe we must hold the textbook industry to a higher standard when it comes to their text development process. We must insist that they provide students with quality texts that, while promoting knowledge, also promote a sense of the inquiry process and sense of joy that guides scientific discovery. This dissertation study is just a starting point. I plan to expand this line of inquiry over the next few years, and I hope others will join the effort. Our students deserve nothing less.

References

American Association for the Advancement of Science (2002). *Middle grades science textbooks: A benchmarks-based evaluation*. Retrieved from http://www.project2061.org/publications/textbook/mgsci/report/about.htm.

Bartlett, F.C. (1932). *Remembering: A study in experimental and social psychology.* Cambridge: Cambridge University Press.

Abrami, G. (1972). Optimum mean temperature for plant growth calculated by a new method of summation. *Ecology* 53, 893-900.

Ackerson, V.L. and Abd-El-Khalick, F.S. (2005). "How should I know what scientists do? I'm just a kid": Fourth-grade students' conceptions of nature of science. *Journal of Elmentary Science Education*, 17(1), 1-11.

Afflerbach, P, and Johnston, P. (1984). On the use of verbal reports in reading research. *Journal of Reading Behavior*, 16 (4), 307-322.

Akinwande, D. ([7 paras.] eHow [online reference]. Available http://www.ehow.com/about_5449400_history-fertilizers.html Spetember 23, 2009.

Allchin, D. (2003). Scientific myth-concpetions. *Science Education,* 87, 329-351.

Altbach, P. G. (1991). Textbooks: The International Dimension. In Apple, W. Michael W. and L.K. Christian-Smith (Eds.), *The politics of the textbook* (pp. 242-249). NY: Routledge.

Anderson, R.T. (1971). Voluntary associations in history. *American Anthropologist*, 7 (3), 209-222.

Apple, M. (1991) Regulating the text: The socio-historical roots of state control. In G.P. Altbach, H.G. Kelly, and L.W. Petrie (Eds.), *Textbooks in American society: Politics, policy and pedagogy* (pp.32-51). New York: State University of New York.

Aronson, E. and Farr, R. (1988). Issues in assessment. *Journal of Reading*, 32 (2), 174-177.

August, S., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research & Practice*, 20(1), 50-57.

Baram-Tsabari, A. and Yarden, A. (2005). Text genre as a factor in the formation of scientific literacy, *Journal of Research in Science Teaching*, 42(4), 403–428.

Bauer, H.H. (1994). *Scientific literacy and the myth of the scientific method*. Champaign-Urbana, IL: University of Illinois Press.

Berman, RA (1995). Narrative competence and storytelling performance. *Journal of Narrative and Life History*, 5, 285-313.

Beveridge, W. I. B. (1950). *The art of scientific investigation.* London: Heinemann.

Biber, Douglas (1992). On the complexity of discourse complexity: A multidimensional analysis. *Discourse Processes*, 15(2) 133-163.

Birkeland, K. (1913). The Norwegian aurora polaris expedition 1902-1903: Vol. I, On the cause of magnetic storms and the origin of terrestrial magnetism (Unknown Binding).http://ia340919.us.archive.org/2/items/norwegianaurorap01chririch/norwegianaurorap01chririch.pdf

Black, J.B. and Wilensky, R. (1979). An evaluation of story grammars. *Cognitive science*, 3, 213-229.
Bowey, J.A. (1986). Syntactic awareness in relation to reading skill and ongoing reading comprehension monitoring. *Journal of Experimental Child Psychology,* 41, 282-299.
Bowker, G. C. (2005). *Memory practices in the sciences*. Cambridge: MIT Press.
Bransford, J. (1985). Schema activation and schema acquisition. In H. Singer & R. B. Ruddell (Eds.), Theoretical models and processes of reading, 3rd ed. Newark, DE: *International Reading Association*, 385-397.
Bruner, J. (1990). *Acts of Meaning*. Cambridge: Harvard University Press.
Bruner, J. (1991). The narrative construction of reality. *Critical Inquiry,* 18(1), 1-21.
California State Board of Education (1998). *Science content standards for California public schools*. Retieved from http://www.cde.ca.gov/be/st/ss/documents/sciencestnd.pdf.
Carey, S. (1985). Are children fundamentally different thinkers than adults? In S. Chipman, J. Segal & R. Glaser (Eds.), *Thinking and learning skills* (pp. 436-517). Hillsdale, NJ: Lawrence Erlbaum.
Cartwright, Nancy (1983): *How the laws of physics lie.* Oxford, New York: Clarendon Press.
Carver, G.W. (1911). Personal correspondence with Booker T. Washington. In J.R. Kremer (1987), *George Washington Carver: In his own words*. Columbia, MS: University of Missouri Press.
Caswell, L.J., & Duke, N.K. (1998). Non-narrative as a catalyst for literacy development. *Language Arts,* 75,108-117.
Clason, D.L. and Dormody, T. J. (1994) Analyzing data measured by individual Likert-type items. *Journal of agricultural education,* 35(4), 31-35.
Cleland, E.E. (2006). Diverse responses of phenology to global changes in a grassland ecosystem. *PNAS,* 103(37), 13740-13744.
Cleland, E.E., Chiarello, N.R., Loarie, S.R., Mooney, H.A. and Field, C.B. (2006). Diverse responses of phenology to global changes in a grassland ecosystem. PNAS, 103 (37), 13740-13744.
Converse, J., and Presser, S. (1986). *Survey questions*. Beverly Hills, CA: Sage Publications.
Coxhead, A. (2000) A New Academic Word List. TESOL Quarterly, 34(2
Crick, F. (1988). *What mad pursuit: A personal view of scientific discovery*. NY: Basic Books.
Cunningham, A. E. & Stanovich, K. E. (1998). What reading does for the mind. *American Educator*, *22*(1&2), 8-15.
Curie, M. (1904). *Radioactive substances*. London: Chemical News Office.
Dale, E. and Chall, J. S. (1948) A formula for predicting readability. *Educational Research Bulletin*, 27. 37-54.
de Castell, S., Luke, A. and Luke, C. (1989). *Language, authority and criticism: Readings on the school textbook*. London: The Falmer Press.
Donaldson, J. and Scheffler, A. (2003). *The snail and the whale*. London: Macmillan.
Donovan, C.A., & Smolkin, L.B. (2001). Genre and other factors influencing teachers' book selections for science. *Reading Research Quarterly*, 36, 412-440.

Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly*, 38 (1), 78-103.
Duke, N. (2000). 3.6 minutes per day: The scarcity of informational texts in first grade. *Reading Research Quarterly*, 35, 202-224.
Dyson, A. H. and Genishi, C. (1994). *The Need for Story: Cultural Diversity in Classroom and Community*. Urbana, IL: National Council of Teachers of English.
Ebbers, M. (2000). Science text sets: Using various genres to promote literacy and inquiry. *Language Arts,* 80(1), 40-50.
Eccles, J. S., Wigfield, A., & Schiefele, U. (1998). Motivation to succeed. In W. Damon (Series Ed.) & N. Eisenberg (Volume Ed.), *Handbook of child psychology* (5th ed., Vol. III). New York: Wiley.
Educational Testing Service (2008). California standards tests (CSTs) technical report, Spring 2007 administration.
Falk, H., & Yarden, A. (2009). "Here the scientists explain what I said." Coordination practices elicited during the enactment of the results and discussion sections of adapted primary literature. *Research in Science Education,* 39(3), 349-383.
Faraday, Michael (1859). *Experimental Researches in Chemistry and Physics* (pp. 33–53). London: Richard Taylor and William Francis.
Feynman, R. (1985). *Surely you're joking, Mr. Feynman!: Adventures of a curious character, with contributions by Ralph Leighton.* NY: W. W. Norton & Co,
Finlay y Barrés , C.E. (1881). Original correspondence. In M.C. Khan (Ed.), *Carlos Finlay and yellow fever*. NY: Oxford University Press: 1940.
Fivush, R. (1994). Constructing narrative, emotion, and self in parent-child conversations about the past. In U. Neisser & R. Fivush (Eds.), *The remembering self: Construction and accuracy of the life narrative* (pp. 136-157). NY: Cambridge University Press.
Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology, 32*, 221–23.
Forster, E.M. (1927). *Aspects of the novel*. London: Penguin books.
French, R.M. (under review). Discovery and creation: Opposite ends of a continuum of constraints. *Philosophical Psychology.* Retrieved from http://docs.google.com/viewer?a=v&q=cache:xloCeMvWhSkJ:www.u-bourgogne.fr/LEAD/people/french/creativity.pdf.
Galilei, G. (1610). *Starry messenger* (*sidereus nuncius)*. translated by A. van Helden, Chicago: 1989.
Gallagher, J.J. (1991) Perspective and practicing secondary school science teachers' knowledge and beliefs about the philosophy of science. *Science Education,* 75, 121-134.
Giddens, A. (1974). *Positivism and sociology*. London: Heinemann.
Giere, Ronald N. (1988). *Explaining science: A cognitive approach.* Chicago: University of Chicago Press.
Glynn, S. M., & Muth, K. D. (1994). Reading and writing to learn science: Achieving scientific literacy. *Journal of Research in Science Teaching*, 31, 1057 -1073.
Goldstein, G. and Hersen, M. (1984). *Handbook of psychological assessment.* NY: Pergamom Press.

Goodall, J. (1971). *In the shadow of man*. NY: Houghton Mifflin.
Gopnik, A. (1996). The scientist as child. *Philosophy of Science,* 63(4), 485-514.
Graesser, A., McNamara, D.S., Louwerse, M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavioral Research Methods, Instruments, and Computer*s, 36, 193-202.
Griffiths, A.K. and Barman, C.R. (1995). High school students' views about the nature of science: Results from three countries. *School Science and Mathematics*, 95, 248-356.
Guth, A. (1997). *The inflationary universe: The quest for a new theory of cosmic origins*. Reading, MA: Perseus Books.
Guthrie, J. T., & Wigfield, A. (2000). Engagement and motivation in reading. In. M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Hand-book of reading research* (3rd. ed., pp. 403–422). New York: Longman.
Guthrie, J.T., Wigfield, A., Barbosa, P., Perencevich, K.C., Taboda, A. and Davis, M.H. (2004). Increasin reading comprehension and engagement through concept-oriented reading instruction. *Journal of Educational Psychology*, 96, 403-423.
Guzzetti, B.J., Williams, W.O., Skeels, S.A., & Wu, S.M. (1997). Influence of text structure on learning counterintuitive physics concepts. *Journal of Research in Science Teaching,* 34, 701-719.
Habermas, J. (1972). *Knowledge and human interests.* trans. Jeremy J. Shapiro. London: Heineman.
Halliday, M. A. K. & Hasan, R. (1976). *Cohesion in English*. London : Longman
Hasbrouck, J. E., & Tindal, G. (1991). Curriculum-based oral reading fluency norms for students in grades 2 through 5. *Teaching Exceptional Children,* 24(3), 41-44.
Hofstadter, D. (1985). On the seeming paradox of mechanizing creativity. In D. R. Hofstadter (Ed.) *Metamagical Themas* (pp. 525-546). NY: Basic Books.
Holmyard, E.J. (1990). *Alchemy*. NY: Dover.
Holton, G. (1988). *Thematic origins of scientific thought*. Cambridge: Harvard University Press.
Hudson, J.A. (1990). The emergence of autobiographical memory in mother-child conversation. In R. Fivush & J.A. Hudson (Eds.), Knowing and remembering in young children (pp.166-196). Cambridge: Cambridge University Press.
Hume, D. (1739). *The treatise of human nature.* Retrieved from http://www.class.uidaho.edu/mickelsen/ToC/hume%20treatise%20ToC.htm.
Isaacson, W. (2007). *Einstein: His life and universe.* NY: Simon & Schuster.
Ivanič, R. (1998). *Writing and identity: The discoursal construction of identity in academic writing*. Amsterdam: John Benjamins.
Jabir (cir. 8th cent.). In E.J. Holmyard and R. Russell (Eds.), *The works of Gerber*. West Glacier, MT: Kessinger: 1928.
Jastan, J.G.B.B. (2009). The effects of homosapians-canis lupus familaris manual abdomenal contact on mood regulation. Unpublished manuscript.
Jenner, E. (1798). An inquiry into the causes and effects of the variolæ vaccinæ, or cow-pox. In C.W. Eliot (Ed.), The Harvard classics. NY: P.F. Collier & Son, 1909-14. Bartleby.com, 2001. www.bartleby.com/38/4/. [Date of Printout].

Jetton, T.L. (1994). Information-driven versus story-driven: What children remember when they are reading informational stories. *Reading psychology*, 15(2), 109-130.
Johnson, N.S. and Mandler, J.M. (1980). A tale of two structures: Underlying and surface forms in stories. *Poetics,* 9, 52-86.
Kandel, E.R. (2007). *In search of memory - The emergence of a new science of mind*. NY: WW Norton & Company.
Kennedy, C.A., Wilson, R. and Draney, K. (2005). Construct map [computer program]. University of California, BEAR Center: Berkeley.
Kintsch, W. (1998). *Comprehension*. Cambridge: Cambridge University Press.
Kintsch, W. and van Dijk, T.A. (1983). *Strategies of discourse comprehension*. London: Academic Press.
Klare, G.R. (1984). Readability. In P. D. Pearson (Ed.), *Handbook of Reading Research* (Vol. 1, pp. 681-744). New York: Longman.
Kuhn, T.S. (1962). *The structure of scientific revolutions* Chicago: Univ. of Chicago Press.
Labov, W. & Waletzky, J. (1997/1967). Narrative analysis: Oral versions of personal experience. Journal of Narrative and Life History, 7(1-4), 3-38.
Landau, M. (1991). *Narratives of human evolution*. New Haven: Yale University .
Latour, B. (1987). *Science in action.* Cambridge: Harvard University Press.
Latour, B., & Strum, S. (1986). Human social origins. Please tell us another origin story! *Journal of Biological and Social Structures,* 9,169-187.
Lederman, N.G. (1992). Students' and teachers' conceptions about the nature of science: A review of the research. *Journal of Research in Science Teaching*, 29, 331-359.
Leitch, T.M. (1986). *What stories are: Narrative theory and interpretation*. University Park: Pennsylvania University Press.
Leite, L. (2002). History of science in science education: Development and validation of a checklist for analyzing the historical content of science textbooks. *Science and Education,* Volume 11, 333-359.
Lemke, J. (1990). *Talking science: Language, learning and values.* Westport, CT: Ablex Publishing.
Linde, C. (2001). Narrative and social tacit knowledge. *Journal of knowledge mangement*, 5(2): 160-171.
Madrazo, G. M. (1997). Using trade books to teach and learn science. *Science and Children*, March, 20-21.
Mahon, B. (2003). *The man who changed everything: The life of James Clerk Maxwell.* Chichester: John Wiley and Son
Marrin, A. (2002). *Dr. Jenner and the speckled monster.* NY: Dutton Children's Books.
Marshall, B. (2002). The discovery that Heliobacter pylori, a spiral bacterium, caused pepetic ulcer disease. In B. Marshall (Ed.), *Heliobacter pioneers: Firsthand accounts from the scientists who discovered heliobacters, 1892-1982.* Chapter 15 (165-201). MA: Blackwell Publishing.
Marshall, B.J. and Warren, J.R. (1984). Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *The lancet*, June 16 1984, 1311-1315.

Martin, M. (1991). Science education and moral education. In M.R. Matthews (ed.), *History, philosophy and science teaching: Selected readings.* Toronto: Teachers College Press.
Mayer, D.A. (1995). How can we best use children's literature in teaching science concepts? *Science and Children* 32, 16-19, 43.
Mayer, R.E., Bove, W., Bryman, A., Mars, R. & Tapangco, L. (1996). When less is more: Meaningful learning from visual and verbal summaries of science textbook lessons. Journal of Educational Psychology, 88(1), 64-73.
McCabe, A. and Peterson, C. (1991). Getting the story: A longitudinal study of parental styles of eliciting narrative and the development of narrative skill. In A. McCabe and C. Peterson (Eds.), *Developing narrative structure* (pp. 217-253). Hillsdale, NJ: Lawrence Erlbaum Associates.
McNamara, D.S., Kintsch, E., Songer, N.B., & Kintsch, W. (1996). Are good texts always better? Text coherence, background knowledge, and levels of understanding in learning from text. *Cognition & Instruction, 14,* 1-43.
McNamara, D.S., Louwerse, M.M., Cai, Z., & Graesser, A. (2005, January 1). Coh-Metrix version 1.4. Retrieved [January 20, 2010], from http//:cohmetrix.memphis.edu.
Medawar, P.B. (1988). *The limits of science*. Oxford: Oxford University Press.
Miller, G.A, Beckwith, R., Fellbaum, C., Gross, D. & Miller, K. (1990). Five papers on WordNet. Cognitive Science Laboratory, Princeton University , No. 43.
Mises, R. (1951). *Positivism: A study in human understanding*. Cambridge, MA: Harvard University Press.
Mishler, E.G. (1995). Models of narrative analysis: A typology. *Journal of narrative life history*, 5(2), 87-123.
Morson, G. S. (1994). *Narrative and freedom: The shadow of time*. New Haven: Yale University Press.
Moss, D.M. (2001). Examinging student conceptions of the nature of science. *International Journal of Science Education*, 23, 771-790.
Nagy, W.E., & Scott, J.A.(2000).Vocabulary processes. In M.L. Kamil, P.B. Mosenthal, P.D. Pearson, & R. Barr (Eds.), *Handbook of Reading Research* (Vol. III, pp. 269–284). Mahwah, NJ: LEA.
Nation, P. (1990). Teaching and learning vocabulary. Boston: Heinle & Heinle.
Nation, P. (2001) *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
National Committee on Science Education Standards and Assessment and National Research Council (1996). *National science education standards.* Washington, DC: National Academy Press. Retrieved from http://www.nap.edu/openbook.php?isbn=0309053269&page=160.
National Institute of Child Health and Human Development (NICHD) (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.
Negrete, A. (2003). Fact via fiction. The pantaneto forum, 12, University of York: http://www.pantaneto.co.uk/issue12/negrete.htm.

Nersessian, N.J. (1999). Model-based reasoning in conceptual change. In L. Magnani, N. J. Nersessian, and P. Thagard (Eds.) *Scientific Discovery* (pp. 5-22). NY: Plenum.
Niaz, M. & Rodriguez, M.A. (2000). Teaching chemistry as rhetoric of conclusions or heuristic principles—A history and philosophy of science perspective. *Chemistry education: Research and practice in Europe*, 1(3) 315-322.
Norris, S. & Phillips, L. (1987). Explanations of reading comprehension: Schema theory and critical thinking theory. Teachers College Record, 89, 2, 281-306.
Ochs, E. (1994). Stories that step into the future. In D.F. Biber and E. Finngan (Eds.), *Sociolinguistic perspectives on register* (pp.106-135). Oxford: Oxford University Press.
Ochs, E. & C. Taylor (1992). Science at Dinner. In C. Kramsch & S. McConnell-Ginet (eds.), *Text and context: Cross-disciplinary perspectives on language study* (pp. 29-45). Lexington, Mass: D.C. Heath.
Ochs, E. & Jacoby, S. (1997). Down to the wire: The cultural clock of physicists and the discourse of consensus. *Language in society*, 479-506.
Ochs, E., Smith, R., & Taylor, C. (1989). Detective stories at dinnertime: Problem-solving through co-narration. *Cultural dynamics*, 2, 238-57.
Orr, J. (1990). Sharing knowledge, celebrating identity: War stories and community memory among service technicians. In D.S. Middleton & D. Edwards (Eds.), *Collective remembering* (pp. 169-189). London: Sage.
Palinscar,A.S. and Magnusson, S.J.(2001)The interplay of first-hand and second-hand investigations to model and support the development of scientific knowledge and reasoning. In S.Carver and D.Klahr (eds) *Cognition and instruction: Twenty-five years of progress* (151-193). Mahwah, NJ: Lawrence Erlbaum Associates.
Pappas, C.C. (2006). The information book genre: It's role in integrated science literacy research and practice. *Reading Research Quarterly,* 41(2), 226-250.
Pearson, P.D., & Camperell, K. (1981). Comprehension of text structures. In J.T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 448-468). Newark, DE: International Reading Association.
Polanyi, L. (1985). *Telling the American story: A structural and cultural analysis of conversational storytelling*. Norwood, NJ: Ablex.
Polya, G. (1945). *How to solve it: A new aspect of mathematical method*. Garden City, NY: Doubleday.
Popper, K. (1934), *The logic of scientific discovery.* London: Routledge (1992).
Prince, G. (1973). *A grammar of stories*. The Hague: Mouton.
Proctor, C. P., August, D., Carlo, M. S., & Snow, C. (2006). The intriguing role of Spanish language vocabulary knowledge in predicting English reading comprehension. Journal of Educational Psychology, 98(1), 159–169.
Propp, V.L. (1928). *Morphology of the Folktale.* Transl.by L Scott, 1958. Austin: Univ. Texas Press.
RAND Reading Study Group. (2002). *Reading for understanding: Towards an R&D program in reading comprehension.*
Rasch, G. (1960/1980). *Probabilistic models for some intelligence and attainment tests.*(Copenhagen, Danish Institute for Educational Research), expanded

edition (1980) with foreword and afterword by B.D. Wright. Chicago: The University of Chicago Press.
Rice, D. C. (2002). Using trade books in teaching elementary science: Facts and fallacies. *The Reading Teacher*, 55(6), 552-565.
Ricoeur, P. (1980). Narrative time. *Critical inquiry*, 7(1), 169-190.
Rosenblatt, L.M. (1989). Writing and reading: The transactional theory. In J. Mason (Ed.), *Reading and writing connections* (pp. 153-175). Boston: Allyn & Bacon.
Rosenblum, Y. & Markovits, A. (1976). Scientific prose or narrative style in a science program for the culturally disadvantaged. Studies in Educational Evaluation, 2, 53–56.
Rumelhart, D.E. (1977a). Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and performance IV* (pp. 573-603). Hillsdale, NJ: Lawrence Erlbaum Associates.
Samuels, S.J. (1977).Introduction to theoretical models of reading. In W. Otto, C. Peters, & N. Peters (Eds.), *Reading problems: A multidisciplinary perspective* (pp.7–41). Reading, MA: Addison-Wesley.
Siegel, H. (1978) Kuhn and Schwab on science texts and the goals of science education. *Educational Theory,* 28:4, 302-309.
Smagorinsky, P. (2001). If meaning is constructed, what is it made from? Toward a cultural theory of reading, *Review of Educational Research, 71 (1*), 135-159.
Snow, C. P. (1963). *A second look: An expanded version of 'the two cultures and the scientific revolution*.' Cambridge: Cambridge University Press.
Sohmer, R., & Michaels, S. (2005). The "Two Puppies" story: The role of narrative in teaching and learning science. In, Quasthoff, U. & Becker, T. (Eds.), *Narrative Interaction* (pp. 57-91). Philadelphia: John Benjamins Publishing.
Spak, K. (2009). Human hair grows plants, profits for Fla. company. [2 paras.] *Newser* [online serial], *5*. Available http://www.newser.com/story/57550/human-hair-grows-plants-profits-for-fla-company.html April 29, 2009.
Sperry, D.E., & Sperry, L.L. (1996). Early development of narrative skills. *Cognitive Development,* 11, 443-465.
Stahl, S. (2003) Vocabulary and readability: How knowing word meanings affects comprehension. *Topics in Language Disorders,* 23(3), 241-47.
Steele C.M. and Aronson, J. (November 1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personal Social Psychology* 69 (5): 797–811
Tasdelen, U. and Koseoglu, F. (2008). Learner-friendly textbooks: Chemistry texts based on a constructivist view of learning. *Asia Pacific Education Review,* 9(2), 136-147.
Thorndyke, P.W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive psychology*, 9(1), 77-110.
Unsworth, L. (1997). "Sound" explanations in school sciences: A functional linguistic perspective on effective apprenticing texts. *Linguistics and Education,* 9(2), 199-226.
Vacca, R. and Vacca, J. (1998). *Content area reading: Literacy and learning across the curriculum*. NY: Addison-Wesley Publishing Co.

van den Broek, P., & Trabasso, T. (1986; Causal networks versus goal hierarchies in summarizing texts. Discourse Processes, 9, 1-15.

van Dijk, T. A., and W. Kintsch. 1983. *Strategies of discourse comprehension*. NY: Academic Press.

Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Chen, R., Pratt, A. and Denckla, M. B. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology*, 88(4) 601–638.

Wellington, J. and Osborne, J. (2001). *Language and literacy in science education*. Philadelphia, PA: Open University Press.

Westerlund, J., & Fairbanks, D., (2004). Gregor Mendel and "myth-conceptions". *Science Education*, 88 (5), 754-758.

White, H. (1980). The value of narrativity. In W.J.T Mitchell (ed.), *On narrative* (pp. 1-24). Chicago: University of Chicago Press.

Wigfield, A., Eccles, J.S., & Pintrich, P.R. (1996). Development between the ages of 11 and 25. In D.C. Berliner & R.C. Calfee (Eds.). *Handbook of educational psychology* (pp. 148-185). New York: Simon & Schuster/Macmillan.

Wilson, M. (2005). *Constructing measures: An item-response modeling approach.* Mahwah, NJ, Lawrence Erlbaum.

Worthy, J., Moorman, M., & Turner, M.(1999). What Johnny likes to readis hard to find in school. *Reading Research Quarterly*, 34 (1), 12-27.

Wu, M.L., Adams, R.J., Wilson, M.R, & Haldane, S.J. (2007). ACERConQuest version 2.0 [computer program]. Hawthorn, Australia: ACER.

Zeno, S.M., Ivens, S.H., Millard, R.T., & Duvvuri, R. (1995). *The Educator's Word Frequency Guide*. Brewster: Touchstone Applied Science Associates, Inc.

Appendices

Appendix A: Experimental Texts

***Radioactivity*, CKE Version (RAD-CKE)**

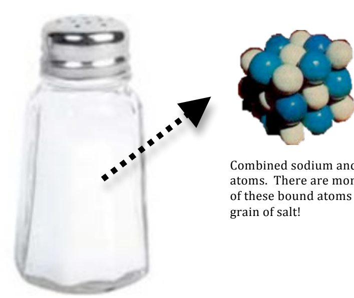
[Page 1]

What are the basic elements that make up the earth? If you could break down a rock into its smallest parts, what would you find? Gold? Copper? Elements are individual pieces of matter that combine with each other and make up everything we see around us. Most of what we see and use in the world, like air and water, is not made of one single element. For example, sodium and chlorine are two different elements that make up the salt that we use for cooking.

We have salt because the elements sodium and chlorine chemically "react" when they are in contact with each other. Chemical reactions occur when separate elements bind together to create a new substance, like salt. The smallest parts (atoms) within these different elements bind together during the chemical reaction.

[Page 2]

We now know of 92 elements that are a natural part of the Earth. By the late 1800s, scientists had already found most of these elements. Some of these elements are more rare than others. Radium is a rare element that was first discovered by Polish-born Marie Curie and her French husband, Pierre, during the late 1800's. Radium is often found in a rock called pitchblende, which also has a large amount of the element called uranium. Pitchblende is so rare that it can only be found in the hills and mountains of a few different countries.



Combined sodium and chloride atoms. There are more than trillions of these bound atoms in a single grain of salt!

[Page 3]

This pitchblende rock was first discovered by Henri Becquerel, who worked with Marie and Pierre Curie in France. Just like any other rock, pitchblende is made of different kinds of elements. What makes pitchblende rare is the high level of the element called uranium along with two special elements—radium and polonium. The reason these elements are special is because they are radioactive.

pitchblende

[Page 4]

While most elements, like sodium and chlorine, cannot change, elements like radium are breaking down (decaying) into another element. This decaying process happens when an element like radium is slowly losing parts of its atoms, turning it into another element. Radium is slowly decaying into the element called lead.

Radium is very unstable because it is decaying. As an element decays, it creates a light blue glow and electricity in the air. These glowing, electric elements can turn part of a copper plate black. We call the source of this glow and electricity "radioactivity." The radioactivity in radium is so powerful that it can pass through almost anything, even a solid brick wall.

[Page 5]

The radioactive elements in pitchblende rock are powerful enough to withstand most acids, even at the highest temperatures. However, non-radioactive elements like iron, carbon and uranium will react with different types of acids and either turn into a gas and dissolve away or turn into whole solid clumps. On the other hand, radioactive elements will remain unchanged.

[Page 6]

Another way that radioactive elements like radium are powerful is that they create electricity in the air. A machine called an electrometer can measure the radioactivity in the form of electricity. Measuring the amount of electricity in the air is a way to know how strong the radioactivity is—the more electricity is in the air, the greater the radioactivity.

electromete

[Page 7]

When radium is in its natural form as part of the pitchblende rock, the electrometer detects less electricity than when the radium is in its pure form, separate from all other elements. As more non-radioactive elements are removed from the pitchblende, the electrometer detects a greater amount of electricity. The greatest amount of electricity is detected when radium is in its pure form, which looks like a small amount of white powder.

Radium in its pure form is a white powder.

[Page 8]

The electrometer is just one way to see the power of radioactivity. Radioactivity can also have powerful effects on people's lives, and some of these effects are very dangerous.

If people come into contact with radium for a long period of time, they can become very sick and even die from radiation poisoning. First, people's hands may become red and swollen from touching the radium, even in its weaker, pitchblende form. Over several months of being near radioactivity, people's hair will begin to fall out and they cannot eat without vomiting. Over time, they grow so tired and weak that they cannot move. Without any treatment, people with radiation poisoning will die. Radioactive elements like radium clearly should be used with great care.

[Page 9]

Radioactivity has the power to destroy lives. But in very small amounts and with full protection, radioactive elements can actually save lives in different ways. For example, just a very tiny amount of radioactivity is used in x-ray machines that show broken bones and certain types of cancer. The glow from the radioactive element can move through the body and make enough light so that an x-ray machine can take a picture. This way, doctors can see what is hidden inside a body without having to do surgery.

Two x-ray images of a left hand

[Page 10]

Once doctors have found where a cancer is, they can use radium to kill the cancer through radiation therapy. In this type of therapy, doctors direct the radiation to the part of the body that has the cancer and weaken it. Once the part with cancer begins to die, the body has a better chance of healing itself.

Patient receives radiation treatment for cancer that is located near her stomach.

[Page 11]

The power of radioactive elements seem magical at first because of the light blue glow that seems to move through almost anything. But when we get a closer look, we can see that these decaying elements are not magical. Radioactivity is what we get when an unstable element like radium breaks down into another element. The glow that we see from these unstable elements can be dangerous or helpful, depending on how it is used. The power of radioactivity can save lives, which is amazing indeed.

***Radioactivity*, SDN version (RAD-SDN)**

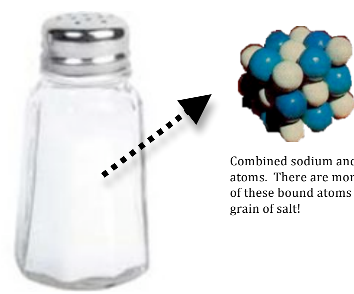
[Page 1]

What are the basic elements that make up the earth? If you could break down a rock into its smallest parts, what would you find? Gold? Copper? Elements are individual pieces of matter that combine with each other and make up everything we see around us. Most of what we see and use in the world, like air and water, is not made of one single element. For example, sodium and chlorine are two different elements that make up the salt that we use for cooking.

We have salt because the elements sodium and chlorine chemically "react" when they are in contact with each other. Chemical reactions occur when separate elements bind together to create a new substance, like salt. The smallest parts (atoms) within these different elements bind together during the chemical reaction.

[Page 2]

We now know of 92 elements that are a natural part of the Earth. By the late 1800s, scientists had already found most of these elements, but some were yet to be discovered. At this time, Polish-born Marie Curie and her French husband, Pierre, were two chemists living in France, trying to find all of Earth's natural elements. Although it was very hard detective work, Marie and Pierre loved solving the mysteries of elements.



Combined sodium and chloride atoms. There are more than trillions of these bound atoms in a single grain of salt!

[Page 3]
One day, a fellow scientist named Henri Becquerel showed Marie and Pierre a special kind of rock called pitchblende. When Henri took this pitchblende rock into a dark room, Marie could see that it gave off a light blue glow. Henri then showed his friends how this glow would turn part of a copper plate to black.

pitchblende

Henri explained that the pitchblende had a lot of the element called uranium, and that he believed that the glow came from the uranium.

[Page 4]
Certainly, this was one of the strangest rocks that Marie and Pierre had ever seen, which was why they wanted to learn as much as possible about the mysterious blue glow and whether it came from the uranium.

Marie and Pierre first built a special instrument called an electrometer, which measures the level of electricity in the air. They noticed that the brighter the glow, the greater the amount of electricity near the rock. They realized that there must be something causing both the electricity and the blue glow.

electromete

[Page 5]
Marie decided to name the cause for both the glow and electricity "radioactivity" because it all "radiated" from the rock. We have used this term "radioactivity" ever since.

Once Marie and Pierre learned how to measure the radioactivity, they wanted to find out exactly which elements in the pitchblende were radioactive. Henri believed this radioactivity was coming from the element uranium, but Marie wasn't sure about this. She and Pierre decided to do some experiments with the pitchblende.

Pitchblende rock was not easy to find, but Pierre and Marie found some in the hills of a nearby country. They took one rock and crushed it into very tiny pieces.

[Page 6]
Like true chemists, Marie and Pierre burned the pitchblende at different temperatures and added different kinds of acid to see what would happen. If they burned the rock too quickly or added too much acid, all of the pitchblende would be

gone or destroyed and they would have to start over again. Marie learned from their experiments that the radioactivity stayed even after many other elements were burned away. She also learned that the radioactivity would not react with most acids. Even when she and Pierre added different acids to the pitchblende, the radioactivity was unchanged.

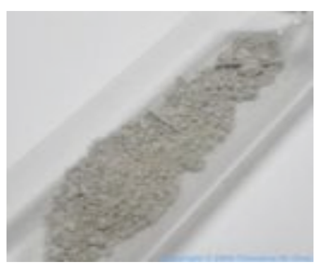
The non-radioactive elements, like iron and carbon, would react with acid and either turn into a gas and dissolve away or turn into whole solid clumps, which could be removed from the pitchblende.

[Page 7]

Marie and Pierre worked for many months to make sure that they were getting rid of all the elements that were not radioactive so that they could see what was left behind. Marie used the electrometer to know whether or not the radioactivity was still in the crushed rock after each experiment. She knew that only non-radioactive elements had been successfully removed if there was as much or even more electricity in the air. When Marie and Pierre finally removed all of the uranium, a high level of radioactivity was still coming from the rest of the rock pieces. This was evidence that the radioactivity did not come from the uranium.

[Page 8]

After a full year of testing the pitchblende, Marie and Pierre finally were able to get rid of every other element except the material with the radioactivity. What was left was a small amount of strange white powder. The electrometer showed that the radioactivity was most powerful when all other elements were removed.



Radium in its pure form
is a white powder.

[Page 9]

Marie and Pierre discovered two new elements in this white powder. The first one she called polonium, after the name of her homeland—Poland—and the second element she named radium because it had strong radioactivity.

Scientists now know that elements like radium are breaking down (decaying) into another element. This decaying process happens when an element like radium is slowly losing parts of its atoms, turning it into another element, like lead. This decaying process makes the radioactive element very unstable and potentially dangerous.

[Page 10]

While working with excitement and hope for discovering a brand new element, Pierre and Marie noticed that they were beginning to feel tired and sick.

They both saw how their hands began to swell and turn red. Pierre's hair was also starting to fall out, and he thought it might have something to do with the radioactivity. He insisted on being the only one to touch the pitchblende, but it was of no use—even a solid brick wall would not have been enough protection. Marie was beginning to feel the same as Pierre, very tired and unable to eat without vomiting. They had not yet realized that the radioactivity was slowly killing them both.

It was too late to stop the effects of radiation on Marie and Pierre. Without protection, both chemists had become poisoned by the radioactivity and remained sick for the rest of their lives. But their hard work and sacrifices helped us to learn about the effects of radiation on people.

[Page 11]

Since they were the first to discover radioactive elements, Marie and Pierre had no idea that radioactivity would someday be used to save lives in different ways. For example, just a very tiny amount of radioactivity is used in x-ray machines that show broken bones and certain types of cancer. The glow from the radioactive element can move through the body and make enough light for an x-ray machine to take a picture. This way, we can see what is hidden inside a body without having to do surgery.

Two x-ray images of a left hand

[Page 12]

Also, doctors use radium to kill the cancer through radiation therapy. In this type of therapy, doctors direct the radiation to the part of the body that has the cancer and weaken it. Once the part with cancer begins to die, the body has a better chance of healing itself.

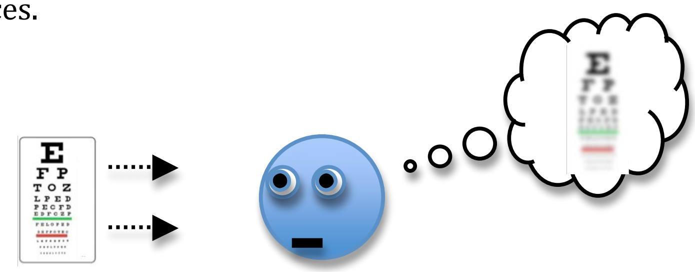
Patient receives radiation treatment for cancer that is located near her stomach.

Marie and Pierre Curie devoted their lives to understanding radioactive elements, and we have all benefitted from their amazing discovery. Although these glowing elements are dangerous, we have learned to use them with great care. Thus, radioactivity has saved many lives.

***Seeing at a Distance*, CKE version (TEL-CKE)**

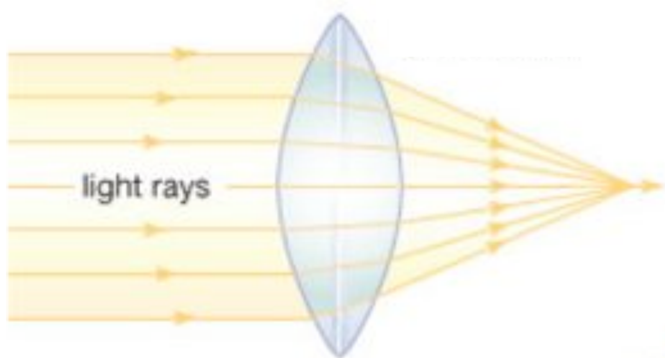
[Page 1]

How far can you see into the distance? Can you see objects from ten feet away, or a hundred feet away? Can you see objects more clearly when they are closer to you? We can see an object because light is travelling to our eyes. In addition to needing light to see around us, some people need glasses to see clearly at different distances.



[Page 2]

Eyeglasses help many people to see clearly because the lenses bend the light as it travels to our eyes. If we look at a side-view of a glass lens, and follow a beam of light as it enters this lens, we would see the light bend. The light bends because the glass "refracts" the light beam. When a glass lens refracts light, we see things differently. Lenses are shaped to refract light in a way that helps those who need glasses see more clearly.



Side view of a convex lens. Light refracts as it enters the lens.

[Page 3]

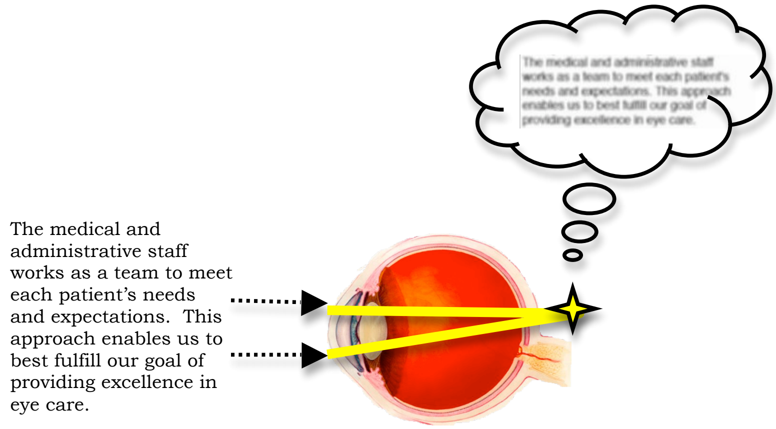
People have been using eyeglasses for seeing more clearly for over a thousand years. But eyeglasses cannot help us see clearly into the far distance, like into the night sky. When we look up into the night sky with our own eyes, what do we see? The moon may look like a smooth ball of light that moves across the sky, but is it? We may think we see many stars on a clear night, but some of those stars are actually planets, like Jupiter.

The telescope is a special tool with eyeglass lenses that can help us see more clearly than we can with our own eyes. With the telescope, we can see Jupiter more clearly—and even its four major moons.

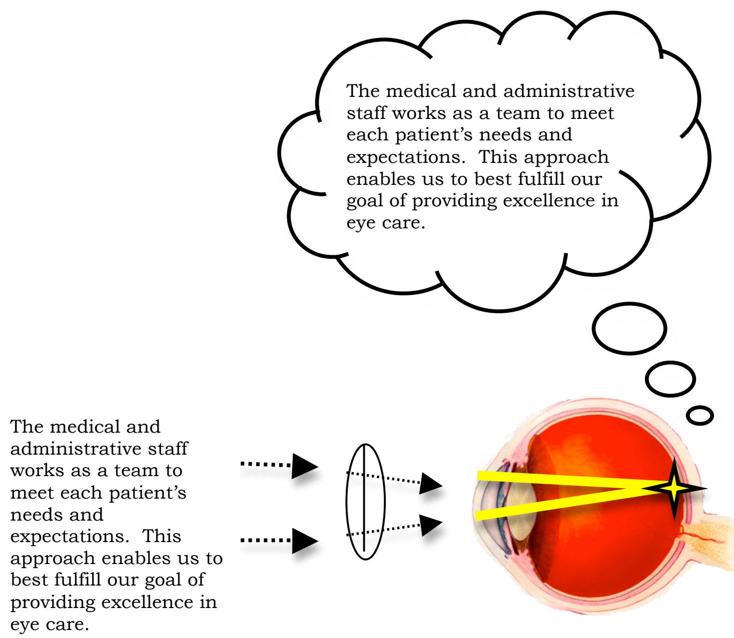
[Page 4]

The word "telescope" comes from the Greek word, "tele-scopos" meaning "far see." A Dutch scientist, Hans Lippershey, first created this special tool in the 1500's. During this time, the famous Italian scientist Galileo created the first telescope that could be used to see into space. The Galilean telescope, which is one of the many types used today, is a tube with two round glass lenses near each end. One of these lenses is convex in shape, meaning that the shape is curved outward. Convex lenses help far-sighted people who cannot clearly see objects that are nearby. The convex curve refracts light "inward" so that the point of focused light is directly on the back of our eyes—the retina. Convex lenses help far-sighted people to focus on objects that are up close, like the words on a page.

[Page 5]



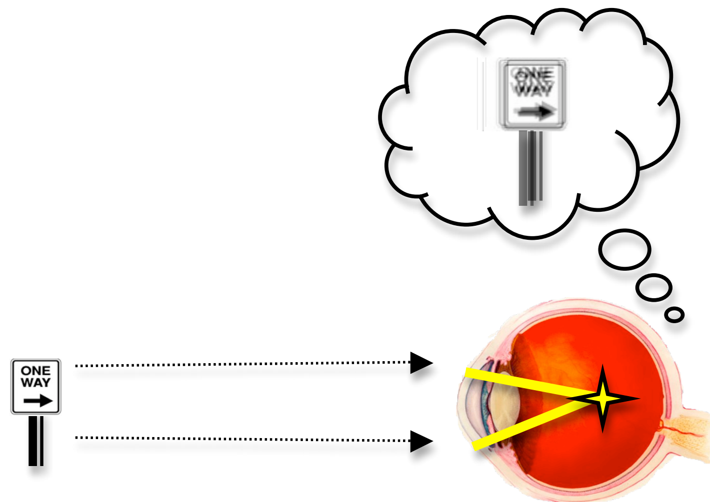
A: Point of focus behind the retina. The words look blurrier than they should.



B: Looking through a convex lens. The lens refracts light inward and the words are now in focus.

[Page 6]

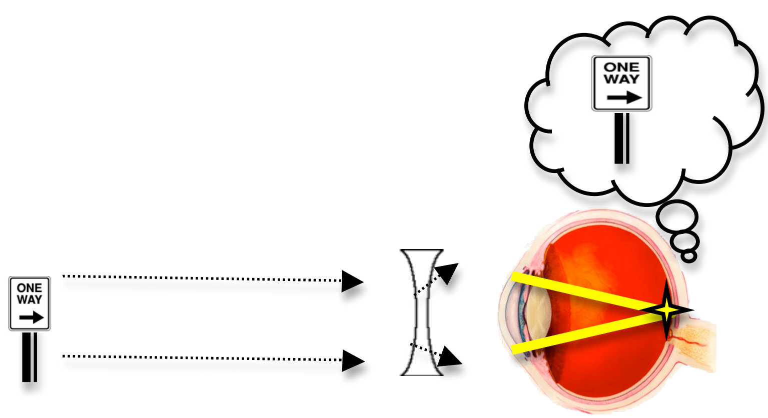
The other type of lens in this telescope is concave in shape, meaning that the front and back are curved inward. Concave lenses help near-sighted people who cannot clearly see objects at a distance. The concave curve of the lens refracts light "outward" so that near-sighted people can focus on distant objects, like a street sign.



A: Point of focus is before the retina, making distant objects unclear.

[Page 7]

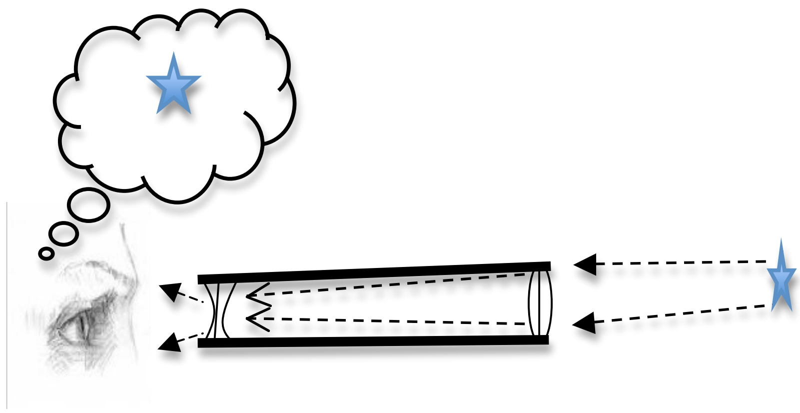
Many telescopes have both convex and concave lenses, which refract light in different ways, helping us to see things at a distance more clearly. For the Galilean telescope, the concave lens is nearest to the eye while the convex lens is at the opposite end of the tube. These two lenses work together to help us see ships 60 miles away in the ocean or planets and stars in the night sky.



B: The concave lens refracts light outward, changing the point of focus on the retina.

[Page 8]

Each lens is carefully shaped so it refracts light in the best way possible. The concave lens refracts more light when its shape has a greater inward curve. In order for the telescope to work well, the concave lens has to refract light more strongly than the convex lens. With a stronger concave lens and a weaker convex lens, we can clearly see the surface of the moon.

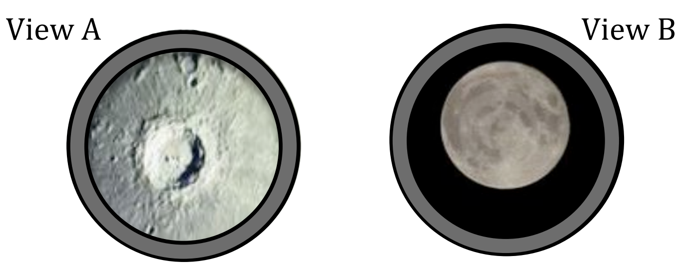


Telescope with concave and convex lenses.

[Page 9]

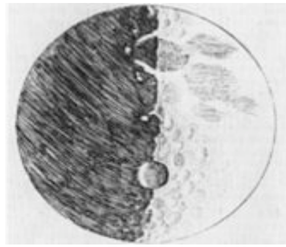
Some telescopes are better than others for seeing details of objects far away. What we see through a telescope is the "field of view." Through the lenses of the Galilean telescope, you can see the mountains and valleys of the moon's surface, but only a section at a time. The tube of this telescope is long enough to capture the details, but the tradeoff for the length is viewing less of the whole object. Because this telescope has a very small field of view, we have to move it around to clearly see all the parts of the moon.

[Page 10]



View A has a smaller field of view than View B.

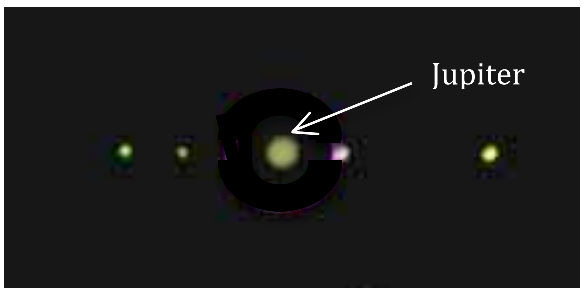
The Galilean telescope is made of only three working parts—two different lenses and a metal tube. And with this simple, powerful tool we can see many details when we use it to look up into the night sky. The moon may look like a smooth ball of light covered with dark spots, but on a closer look through this telescope, we can see deep valleys and great mountain ranges. Through the telescope, we can now see all the different marks on the moon's surface.



Drawing of the moon.

[Page 11]

The moon is only one of the amazing sights to see through our Galilean telescope. We can also see Jupiter, and even though we cannot see the details of Jupiter's surface as we can with the moon, we can see how its moons rotate around it.

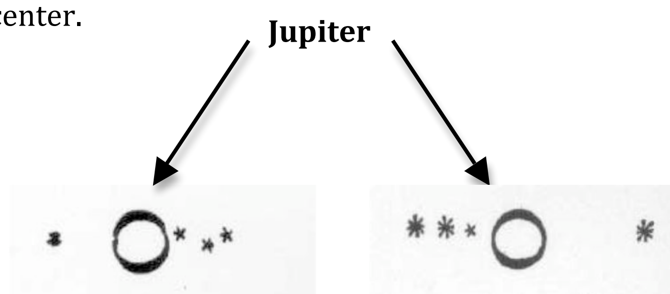


What we can see through the Galilean telescope.

These moons look like bright stars that rotate around Jupiter, always in the same direction, just as our moon rotates around the Earth. If we make different drawings of Jupiter and its moons every night, we can see how the moons move in the same pattern around Jupiter.

[Page 12]

The telescope allows us to learn more about the universe and our place in it. The moons of Jupiter, for example, help us understand that just as our moon rotates around the Earth, there are other moons rotating around other planets. Our solar system contains many planets and moons that all rotate in their individual patterns but are also rotating around the sun. This sun-centered solar system has been called the Copernican model, named after the Polish astronomer, Copernicus, who first claimed that the sun was the center.



Two quick drawings of Jupiter and its moons. These drawings are from two different nights.

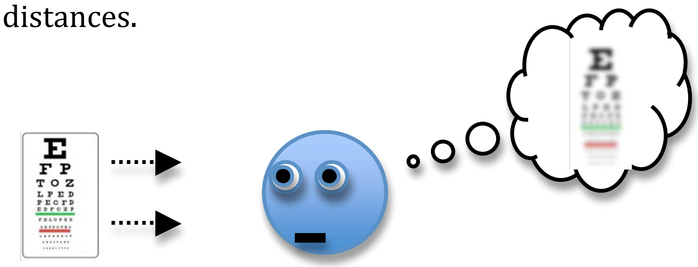
[Page 13]

There is so much to see and learn about the night sky, and the telescope gives us more powerful eyes to learn as much as we can. The lenses of the Galilean telescope have led the way to even better telescopes that help us see even more moons revolving around Jupiter. With our ability to see far into the distance, we can understand the Earth and its place in the universe.

***Seeing at a Distance*, SDN version (TEL-SDN)**

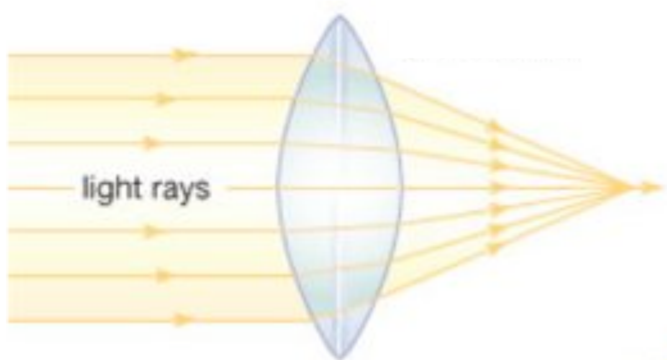
[Page 1]

How far can you see into the distance? Can you see objects from ten feet away, or a hundred feet away? Can you see objects more clearly when they are closer to you? We can see an object because light is travelling to our eyes. In addition to needing light to see around us, some people need glasses to see clearly at different distances.



[Page 2]

Eyeglasses help many people to see clearly because the lenses bend the light as it travels to our eyes. If we look at a side-view of a glass lens, and follow a beam of light as it enters this lens, we would see the light bend. The light bends because the glass "refracts" the light beam. When a glass lens refracts light, we see things differently. Lenses are shaped to refract light in a way that helps those who need glasses see more clearly.



Side view of a convex lens. Light refracts as it enters the lens.

[Page 3]

People have been using eyeglasses for seeing more clearly for over a thousand years. But eyeglasses cannot help us see clearly into the far distance, like into the night sky. Galileo was a scientist during the late 1500's in Italy who looked up into the night sky and wondered, "If I could see everything there is to see, what would I find?"

From far away, the moon seemed like a blotched but smooth ball that moves across the sky. In fact, many people during Galileo's time believed that the moon was a perfectly smooth ball of light. But what is the moon really like?

[Page 4]

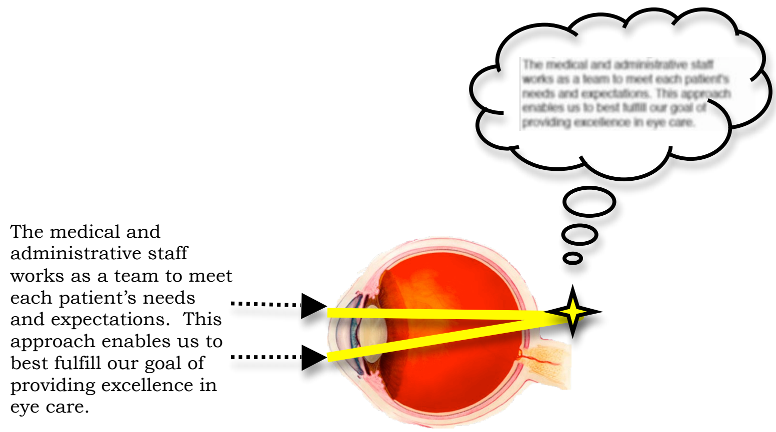
People during this time also believed that the earth was the center of the universe and that everything in space, including the stars and planets, rotated around the earth. Having read the work of Copernicus, Galileo came to wonder whether this was true. Copernicus was a Polish astronomer who argued that the Earth was not the center of the universe, but rather that the stars and planets revolved around the sun.

Galileo wanted to learn as much as he could about the night sky. When Galileo heard about a special tool, a "seeing tube" that would allow him to see objects far away, he knew exactly what he wanted to do. He would make a tube of his own that was powerful enough to study the moon and possibly other planets.

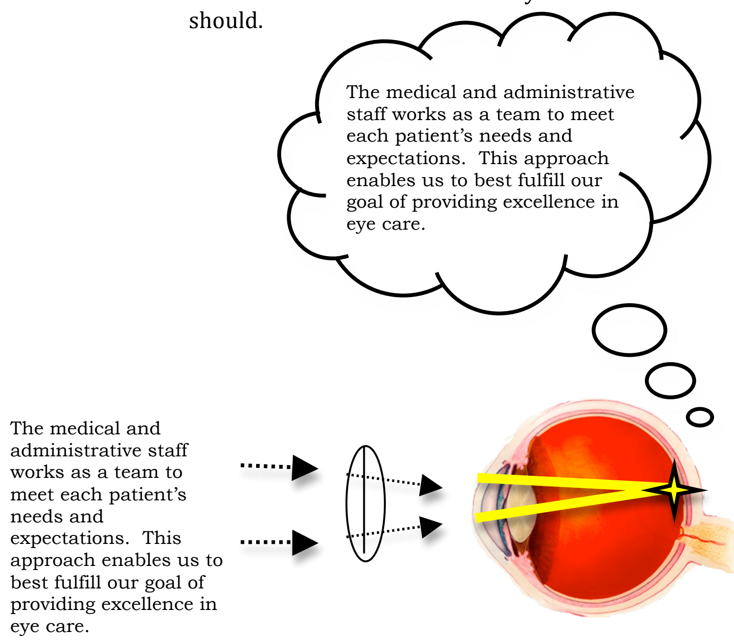
[Page 5]

Galileo learned that a Dutch man named Hans Lippershey had stacked together two lenses from two different types of eyeglasses. One type of lens was convex in shape, meaning that the shape is curved outward. Convex lenses help far-sighted people who cannot clearly see objects that are nearby. The convex curve refracts light "inward" so that the point of focused light is directly on the back of our eyes—the retina. Convex lenses help far-sighted people to focus on objects that are up close, like the words on a page.

[Page 6]



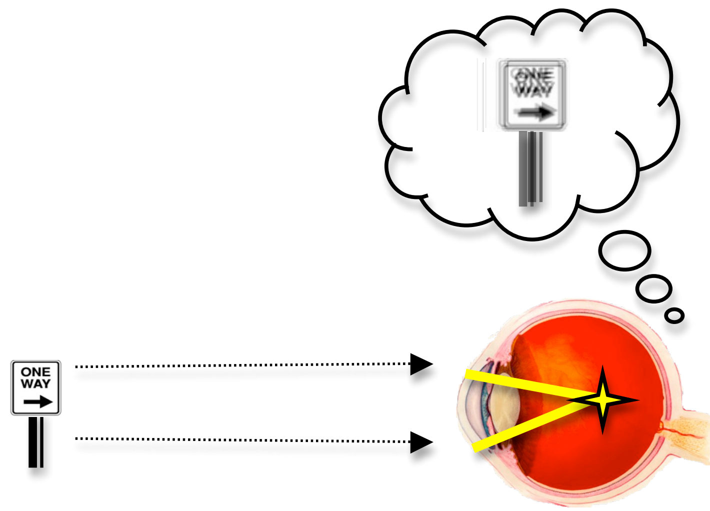
A: Point of focus behind the retina. The words look blurrier than they should.



B: Looking through a convex lens. The lens refracts light inward and the words are now in focus.

[Page 7]

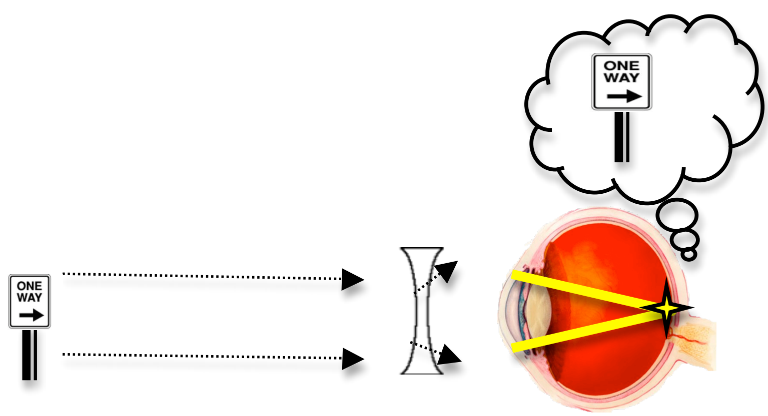
The other type of lens that Lippershey used was concave in shape, meaning that the front and back curved inward. Concave lenses help near-sighted people who cannot clearly see objects at a distance. The concave curve of the lens refracts light "outward" so that near-sighted people can focus on distant objects, like a street sign.



A: Point of focus is before the retina, making distant objects unclear.

[Page 8]

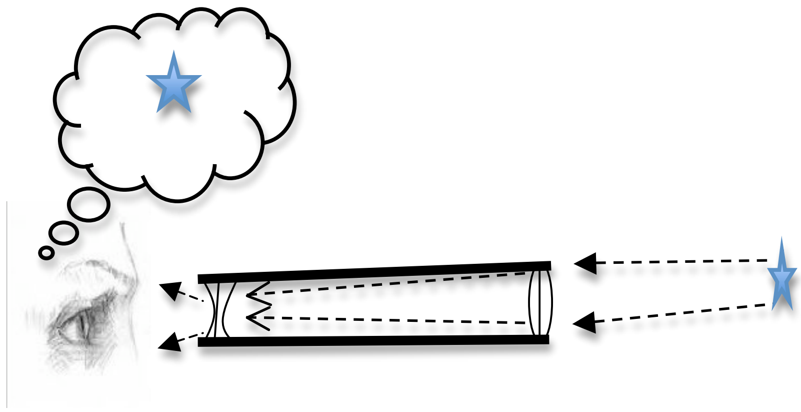
When Lippershey looked through the concave and convex lenses together, objects far away looked like they were closer than they really were. Lippershey built a metal tube to hold these lenses in place. He put the concave lens in the tube nearest to his eye and the convex lens in the opposite end of the same tube. With the lenses in this tube, he could stand on a cliff and clearly see ships on the ocean about 60 miles away. Later on, this tube would be called a telescope. The word "telescope" comes from the Greek word "telescopos," meaning "far see."



B: The concave lens refracts light outward, changing the point of focus on the retina.

[Page 9]

Galileo wanted to make a telescope powerful enough to see clearly into the night sky. He made the tube of his telescope from lead and attached the concave and convex lenses at each end, just like Lippershey. Galileo knew that in order to have lenses powerful enough to see details of objects in space, like the moon, he had to learn how to shape the lenses carefully so that they could refract light in the best way possible.



Telescope with concave and convex lenses.

[Page 10]

To make the best lenses, Galileo first had to find clear glass that had no air bubbles, which was hard to find during his time. He worked all day for many months cutting, grinding and polishing countless glass lenses, carefully shaping each one. Some of his lenses refracted more strongly than others, giving a greater change in what he could see. He took careful notes on how he changed the shape and thickness of each lens before he would try again with a new lens. Most of the lenses that he created did not work, but he learned from making each one. Finally, he learned that the concave lens had to refract light more strongly than the convex lens in order to clearly see the surface of the moon. To make the concave lens refract more, the shape of the lens had to have a greater inward curve.

[Page 11]

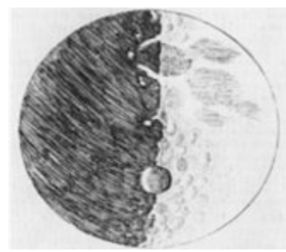
When Galileo looked through his new telescope, he could see the surface of the moon, and so he began his first close look into space. He slept during the day in order to work and see the moon at night. Many people thought that the moon was a smooth ball with a light of its own. Now that Galileo had a closer look through his telescope, he realized that the moon's surface had mountains and valleys.

Even though Galileo could see a lot of detail, his telescope would only let him see a small part of the moon at a time. The longer he made his tube, the more detail he could see, but only in smaller parts. The tradeoff for seeing more detail is that he would see less of the whole moon. What is seen through the telescope is the "field of view." Because his telescope had a very small field of view, he had to move his telescope around to clearly see all the parts of the moon.

[Page 12]

View A has a smaller field of view than View B.

Galileo decided to study the moon every few minutes throughout the night for several months. Each time he looked through his telescope, he would make drawings of every detail. Then, he made a whole drawing from all the parts.



Drawing of the moon.

[Page 13]

During Galileo's time, people believed that everything we could see in the night sky moved around the Earth. Galileo wondered if this was true, and so with his mighty telescope, he began to study the planet called Jupiter. Throughout the night, he would make his drawings of Jupiter, taking note of every mark through his powerful lenses.

What we can see through the Galilean telescope.

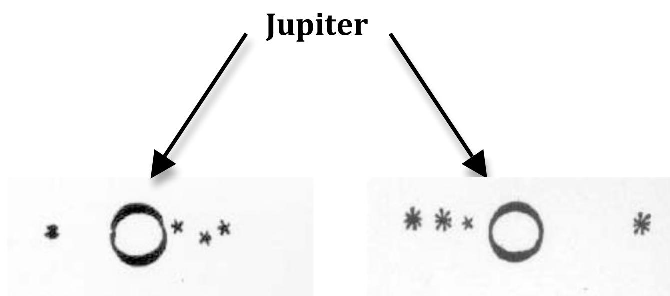
[Page 14]

He looked through his telescope every night, just as he did with the moon. At first, he thought that there were four bright stars near Jupiter, but these stars seemed to move around Jupiter in the same direction every night. After a few months and many drawings, he realized that these stars were actually four moons.

Before this time, no one knew that there were other moons out in space. Galileo learned from his drawings that these moons rotate around Jupiter just as our moon rotates around the Earth. Galileo was the first to have evidence that the Earth was not at the center of everything, as everyone had believed. Although his drawings did not prove that the sun was the center of everything as Copernicus argued, discovering the moons of Jupiter was another step towards the truth about our universe.

[Page 15]

Galileo was excited to discover the mountains and valleys of our moon and the four major moons of Jupiter. The lenses of Galileo's telescope have led the way to even better telescopes that can help us see even more moons revolving around Jupiter. These discoveries about our solar system help us to understand the Earth and its place in the universe.



Two quick drawings of Jupiter and its moons.
These drawings are from two different nights.

Appendix B: Co-Metrix Output for All Four Experimental Texts

The following tables show indices organized by six categories: a) general identification, b) readability, c) general word and text information, d) syntactic quality, e) referential and semantic quality, and f) situational model dimensions. Some indices are presented more than once in because they relate to more than one category.

1. General Identification and Reference Information

Title	Description of text	Genre	Source	LSASpace
RAD_CKE	CKE version of *Radioactivity*	Science	Self generated	CollegeLevel
RAD_SDN	SDN version of *Radioactivity*	Science	Self generated	CollegeLevel
TEL_CKE	CKE version of *Seeing at a Distance*	Science	Self generated	CollegeLevel
TEL_SDN	SDN version of *Seeing at a Distance*	Science	Self generated	CollegeLevel

2. Readability Indices

Title	READASL 'Average Words per Sentence'	READASW 'Average Syllables per Word'	READFRE 'Flesch Reading Ease Score (0-100)'	READFKGL 'Flesch-Kincaid Grade Level (0-12)'
RAD_CKE	16.807	1.62	52.724	10.081
RAD_SDN	18.104	1.579	54.876	10.103
TEL_CKE	18.938	1.363	72.303	7.879
TEL_SDN	18.761	1.375	71.468	7.952

3. General Word and Text Information

3.1. Basic count

Title	READNP 'Number of Paragraphs'	READNS 'Number of Sentences'	READNW 'Number of Words'	READAPL 'Average Sentences per Paragraph'	READASW 'Average Syllables per Word'	READASL 'Average Words per Sentence'
RAD_CKE	14	57	958	4.071	1.62	16.807
RAD_SDN	21	67	1213	3.19	1.579	18.104
TEL_CKE	14	55	1210	4	1.363	18.938
TEL_SDN	17	71	1332	4.176	1.375	18.761

3.2. Frequencies

Title	FRQCRacw 'Celex, raw, mean for content words (0-1,000,000)'	FRQCLacw 'Celex, logarithm, mean for content words (0-6)'	FRQCRmcs 'Celex, raw, minimum in sentence for content words (0-1,000,000)'	FRQCLmcs 'Celex, logarithm, minimum in sentence for content words (0-6)'
RAD_CKE	2953.038	2.367	27.188	1.057
RAD_SDN	2553.072	2.363	27.067	1.08
TEL_CKE	2102.694	2.414	25.729	1.346
TEL_SDN	2410.161	2.458	29.256	1.392

3.3. Concreteness

Title	WORDCacw 'Concreteness, mean for content words'	WORDCmcs 'Concreteness, minimum in sentence for content words'
RAD_CKE	376.15	186
RAD_SDN	368.516	158
TEL_CKE	420.288	158
TEL_SDN	412.345	158

3.4. Hypernymy

Title	HYNOUNaw 'Mean hypernym values of nouns'	HYVERBaw 'Mean hypernym values of verbs'
RAD_CKE	4.345	1.403
RAD_SDN	4.499	1.352
TEL_CKE	4.96	1.618
TEL_SDN	5	1.626

4. Syntactic Indices

4.1. Constituents

Title	DENSNP 'Noun Phrase Incidence Score (per thousand words)'	SYNNP 'Mean number of modifiers per noun-phrase'	SYNHw 'Mean number of higher level constituents per word'	SYNLE 'Mean number of words before the main verb of main clause in sentences'	DENNEGi 'Number of negations, incidence score'
RAD_CKE	264.092	0.921	0.731	5.439	5.219
RAD_SDN	259.687	0.908	0.732	4.761	8.244
TEL_CKE	276.285	0.875	0.722	4.688	4.715
TEL_SDN	280.781	0.816	0.733	4.648	6.757

4.2. Pronouns, types and tokens

Title	DENSPR2 'Ratio of pronouns to noun phrases'	DENPRPi 'Personal pronoun incidence score'	TYPTOKc 'Type-token ratio for all content words'
RAD_CKE	0.15	39.666	0.494
RAD_SDN	0.19	49.464	0.519
TEL_CKE	0.22	60.82	0.288
TEL_SDN	0.225	63.063	0.404

4.4. Connectives

Total connectives

Title 'Title'	CONi 'Incidence of all connectives'	DENCONDi 'Number of conditional expressions, incidence score'
RAD_CKE	67.85	2.088
RAD_SDN	88.211	2.473
TEL_CKE	54.691	1.886
TEL_SDN	54.805	2.252

Positive connectives.

Title 'Title'	CONADpi 'Incidence of positive additive connectives'	CONTPpi 'Incidence of positive temporal connectives'	CONCSpi 'Incidence of positive causal connectives'	CONLGpi 'Incidence of positive logical connectives'
RAD_CKE	27.14	9.395	25.052	18.789
RAD_SDN	45.342	11.542	23.908	22.259
TEL_CKE	18.388	7.544	20.273	18.388
TEL_SDN	18.769	8.258	20.27	15.766

Negative connectives.

Title 'Title'	CONADni 'Incidence of negative additive connectives'	CONTPni 'Incidence of negative temporal connectives'	CONCSni 'Incidence of negative causal connectives'	CONLGni 'Incidence of negative logical connectives'
RAD_CKE	5.219	0	0	5.219
RAD_SDN	9.893	0	1.649	11.542
TEL_CKE	8.015	0	0.943	8.958
TEL_SDN	6.757	0	1.502	8.258

4.5. Logical Operators

Title 'Title'	DENLOGi 'Logical operator incidence score (and + if + or + cond + neg)'
RAD_CKE	34.447
RAD_SDN	52.762
TEL_CKE	24.045
TEL_SDN	26.276

4.6. Sentence syntax similarity

Title 'Title'	STRUTa 'Sentence syntax similarity, adjacent'	STRUTt 'Sentence syntax similarity, all, across paragraphs'	STRUTp 'Sentence syntax similarity, sentence all, within paragraphs'
RAD_CKE	0.114	0.1	0.098
RAD_SDN	0.1	0.085	0.08
TEL_CKE	0.092	0.086	0.093
TEL_SDN	0.107	0.092	0.103

5. Referential and Semantic Indices

5.1. Anaphor

Title 'Title'	CREFP1u 'Anaphor reference, adjacent, unweighted'	CREFPau 'Anaphor reference, all distances, unweighted'
RAD_CKE	0.268	0.096
RAD_SDN	0.348	0.153
TEL_CKE	0.369	0.183
TEL_SDN	0.443	0.197

5.2. Co-reference

Title 'Title'	CREFA1u 'Argument Overlap, adjacent, unweighted'	CREFAau 'Argument Overlap, all distances, unweighted'	CREFS1u 'Stem Overlap, adjacent, unweighted'	CREFSau 'Stem Overlap, all distances, unweighted'	CREFC1u 'Proportion of content words that overlap between adjacent sentences'
RAD_CKE	0.786	0.571	0.821	0.629	0.166
RAD_SDN	0.712	0.592	0.652	0.574	0.141
TEL_CKE	0.856	0.607	0.82	0.548	0.189
TEL_SDN	0.871	0.64	0.843	0.562	0.19

5.3. Latent Semantic Analysis (LSA)

Title 'Title'	LSAassa 'LSA, Sentence to Sentence, adjacent, mean'	LSApssa 'LSA, sentences, all combinations, mean'	LSAppa 'LSA, Paragraph to Paragraph, mean'
RAD_CKE	0.484	0.39	0.653
RAD_SDN	0.427	0.379	0.649
TEL_CKE	0.525	0.521	0.652
TEL_SDN	0.513	0.508	0.636

6. Causal Dimension

6.1. Causal dimension

Title 'Title'	CAUSVP 'Incidence of causal verbs, links, and particles'	CAUSC 'Ratio of causal particles to causal verbs (cp divided by cv+1)'
RAD_CKE	63.674	0.632
RAD_SDN	67.601	0.596
TEL_CKE	59.877	0.542
TEL_SDN	60.06	0.558

6.2. Intentional dimension

Title 'Title'	INTEC 'Ratio of intentional particles to intentional content'	INTEi 'Incidence of intentional actions, events, and particles.'
RAD_CKE	0	8.351
RAD_SDN	0	10.717
TEL_CKE	0	18.388
TEL_SDN	0	18.769

6.3. Temporal dimension

Title 'Title'	TEMPta 'Mean of tense and aspect repetition scores'
RAD_CKE	0.777
RAD_SDN	0.833
TEL_CKE	0.775
TEL_SDN	0.807

6.4. Spatial dimension

Title 'Title'	SPATC 'Mean of location and motion ratio scores.'
RAD_CKE	0.423
RAD_SDN	0.407
TEL_CKE	0.458
TEL_SDN	0.472

Appendix C: Measures

1. Items for Recall, Phase I for Both Topics.

Radioactivity Text, Recall I	Seeing at a Distance Text, Recall I
1. What did you find most interesting from this text? **(open-ended)**	1. What did you find most interesting from this text? **(open-ended)**
2. Do you have any questions of your own about this topic? Write these questions below. **(open-ended)**	2. Do you have any questions of your own about this topic? Write these questions below. **(open-ended)**
3. Not all of the items below are elements. Circle <u>only</u> the elements (eight substances and elements are listed)	3. Draw and label the parts of the Galilean telescope. (open-ended)
4. Your friend shows you a glow stick and thinks it is radioactive because it glows. Explain at least **<u>two ways</u>** to show your friend that the glow stick is not radioactive. **(open-ended)**	4. Use two of the following words (list) to fill in the blanks in the following sentence: "In order to clearly see details on distant objects, the ____(**concave**)_____ lens must have a greater curve than the ___(**convex**)________ lens in a telescope.
3. Which of the following is an element in pitchblende? a.) rock **b.) polonium** c.) gold d.) chlorine	5. A person will see an object most clearly when the light travels directly to the: a.) object b.) eye **c.) retina** d.) lens
6. Radioactive elements are decaying, which means that: a.) they are slowly dying out b.) they are generally unchanged **c.) they are losing parts of atoms** d.) they are gaining parts of atoms	6. Finish the following statement: "Learning about the moons of Jupiter helped us understand . . . "
7. An electrometer can: a.) only measure glowing objects b.) measure chemical reactions **c.) measure levels of electricity** d.) directly measure radioactivity	7. Circle the picture below with the largest field of view. (series of four pictures from the book, *Zoom*)

8. Which of the following is true about radioactivity? a.) just a little bit of exposure will kill you. b.) it is another form of electricity. **c.) it can move through almost anything.** d.) it is less powerful in pure form.	8. Only one of the following statements is true. Circle <u>only</u> the true statement. **a.) The surface of the moon is similar to the Earth's surface.** b.) The Earth moves around the moon in the same direction. c.) The Earth's moon moves around Jupiter in the same direction. d.) The moon is a smooth ball of light and moves around Earth.
9. Circle <u>only</u> the radioactive elements (list of 8 elements and substances presented, like in item 3 for first recall).	9. Which is better for looking at small details on an object far away (circle one)→ long telescope short telescope Explain your choice below. **(open-ended)**
10. In order to have a chemical reaction, **a.) the atoms of different elements must bind together.** b.) the different elements must be radioactive. c.) the atoms of different elements must dissolve. d.) the different elements must have the same atoms.	10. Which of the following diagrams is the <u>**best**</u> example of refraction? (series of four diagrams presenting a grey box and arrows pointed in different directions)

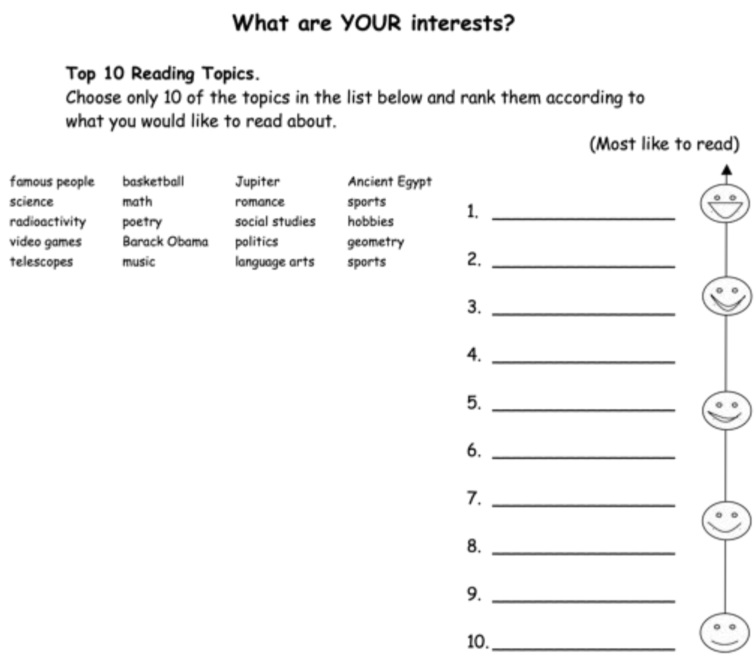
2. Items for Recall, Phase II, for both topics.

Radioactivity Text, Recall II	Seeing at a Distance Text, Recall II
1. Have you tried to learn more about radioactivity since you last read this text? (circle one): YES NO Write what you have done to learn more. **(open-ended)**	1. Have you tried to learn more about radioactivity since you last read this text? (circle one): YES NO Write what you have done to learn more. **(open-ended)**
2. An element: a) always breaks down into smaller elements **b) sometimes breaks down into smaller elements** c) never binds with other elements d) always binds with other elements	2. Light travels: **a) in a straight line directly to our eyes.** b) in a crooked line directly to our eyes. c) in a circle that moves away from our eyes. d) in a straight line directly away from our eyes.
3. An atom: a) is made up of different elements b) is the largest part of matter **c) is the smallest part of matter** d) is the same size as a grain of salt	3. people who need reading glasses generally use: a) concave lenses **b) convex lenses** c) both convex and concave lenses d) thin lenses
4. Radioactivity can save lives by: **a) killing parts of the body with cancer** b) dissolving parts of the body with cancer c) binding with parts of the body with cancer d) making people with cancer stronger	4. refraction is: a) curved light **b) bended light** c) bright light d) bounced light
5. Radioactive elements are different from non-radioactive elements. Write down all the possible ways that make radioactive elements special. **(open-ended)**	5. Match the following two lenses with the correct label: a) concave b) convex

6. Pitchblende is: a) an ordinary rock b) a radioactive element **c) a radioactive rock** d) an uranium element	6. Before Galileo, telescopes could only help us see ships in the ocean 60 miles away. List everything that makes the Galilean telescope much more powerful.
7. Chemical reactions help us understand: **a) what elements bind together** b) what elements are radioactive c) how large atoms are in matter d) how many atoms are in an element	7. The major moons of Jupiter show us that: a) planets move in opposite directions **b) the Earth is not the center of the universe** c) Jupiter is larger than the Earth d) Jupiter is the center of the universe
8. Radioactive elements are dangerous because: a) they have a strange blue glow **b) they break down matter** c) they move through matter d) they are found all around us	8. The Galilean telescope: a) is not the first type to show us details of our moon. **b) cannot show us all the details of the moon at once.** c) can show us all the details of Jupiter. d) is the only telescope used today
9. Circle <u>only</u> the radioactive elements (list of substances and elements).	9. Circle the picture below with the <u>smallest</u> field of view. (Another series of pictures from the book *Zoom*)

3. Interest Measures

For both the initial and subsequent recall sessions, subjects were asked to select 10 topics from the list provided that they would most like to read about and rank them according to their preference. Below is a sample of what each subjects received (list of topics were randomly displayed for each subject).



What are YOUR interests?

Top 10 Reading Topics.
Choose only 10 of the topics in the list below and rank them according to what you would like to read about.

(Most like to read)

famous people	basketball	Jupiter	Ancient Egypt
science	math	romance	sports
radioactivity	poetry	social studies	hobbies
video games	Barack Obama	politics	geometry
telescopes	music	language arts	sports

1. __________
2. __________
3. __________
4. __________
5. __________
6. __________
7. __________
8. __________
9. __________
10. __________

In addition to the interest ranking tasks, there were 4 items interest-related items embedded in the recall tasks described above:

- (For both topics, first item in Recall I task): *What did you find most interesting about this text?*

- (For both topics, second item in Recall I task): *Do you have any questions of your own about this topic? Write these questions below.*

- (Last item after Recall I task): *Which book was most interesting to you? (circle one): telescopes radioactivity Explain your choice.*