This is to certify that the
thesis entitled

THE EFFECT OF COLOR CONTRAST OF TEXT ON THE NOTICEABILITY AND LEGIBILITY OF PRESCRIPTION DRUG WARNING LABELS

presented by

RAGHAV PRASHANT SUNDAR

has been accepted towards fulfillment
of the requirements for the

MS degree in PACKAGING

Major Professor's Signature

7/2/09

Date

THE EFFECT OF COLOR CONTRAST OF TEXT ON THE LEGIBILITY AND NOTICEABILITY OF PRESCRIPTION DRUG LABELS

By

Raghav Prashant Sundar

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Packaging

2009

ABSTRACT

THE EFFECT OF COLOR CONTRAST OF TEXT ON THE LEGIBILITY AND NOTICEABILITY OF PRESCRIPTION DRUG LABELS

By

Raghav Prashant Sundar

Prescription drug warning labels (PWLs), small, colorful, auxiliary stickers applied by pharmacy personnel on prescription vials, are used to provide important, drug-specific information about medications. At present, the required information and the presentation of said labels is not standardized in the US, and they are not regulated by the FDA (Ault 2007). This is despite the fact that, these labels contain information that is critical to the safe and effective use of the products contained within.

Thirty-three subjects from two populations (18-29 and 50+) were characterized in a variety of ways and tested to objectively examine the performance of PWLs in five color contrasts. Tests of label performance included: eye tracking, recall and recognition and legibility. Data were evaluated in various ways.

Statistical analysis was performed using mixed models. There was not enough evidence to conclude an effect of color contrast on the noticeability of the PWL's, although findings indicated that subjects were more likely to notice and spend more time on the large white pharmacy label as compared to the cap and the PWL. It was revealed that subjects from the younger population were more likely to recognize the PWL's they had seen. It was also identified that certain categories of information (dosage, patient information) were recalled more than the others. Legibility findings were consistent with previous studies with the black on white color contrast being the easiest to read.

For my Madhumitha,
You are nothing short of my Everything

Acknowledgements

"… and when you want something, all the universe conspires in helping you achieve it"

- Paulo Coelho in his 1988 book, The Alchemist

Despite its staggering achievements humanity still lacks the ability to contain the universe on this sheet of paper. Nevertheless, I would surely like to bring to light the efforts of several remarkable individuals without whose inputs this thesis would have been impossible to write.

I would like to start off by thanking my advisor, Dr. Laura Bix for her support and guidance throughout my time as a Masters student. It was only due to her supervision and valuable advice that I was able to take up this research project, let alone write this thesis. Dr.Bix is a friend to her students and always encourages us to think "out of the box" and come up with innovative solutions to our problems. It is only through her positive attitude and approach to learning that most of us find our graduate level work fulfilling.

I would also like to thank the other members of my academic committee, Dr.Rafael Auras and Dr.Theresa Mastin, who with their insights, have time and again propelled my work to greater heights.

I would like to thank Mr. Tom Ball from the Oil Health Center who supplied us with all the vials and labels that were used in this study. Mr.Ball was always available to provide us with his inputs on our research from a pharmacy perspective. This was instrumental in setting up many of the experimental methods in this study.

All my fellow students from Dr.Bix's Packaging HUB (Human Factors, Universal Design and Bio Mechanics) Research group were indispensable to the smooth working of this project. The theme for this entire study came off a Plan B project by Voramanus

Viraporn. Without her systematic report and literature review, this study would not have materialized.

My dear friends Chris Steckler and Kirit Rosario helped me design and build the eye tracking chin rest and table. Through this, we were able to achieve a high degree of accuracy using our instruments and also keep older consumers comfortable during testing.

None of the mind-numbing statistics seen in this work would've been possible without Nora Bello, our statistician from the Statistical Consulting Center at CANR. With her suggestive and innovative approach, we were able to make sense of all the data that were had collected.

I would also like to extend my heartfelt thanks to the Center for Food and Pharmaceutical Packaging Research (CFPPR) for their financial support. Additionally, their feedback to me during our periodic meetings was enlightening.

I wouldn't have gotten far in the USA if it weren't for my supportive and understanding employers and co-workers at the Division of Engineering Computing Services, College of Engineering. Working for them as a graduate assistant web developer / graphic designer did a lot more than just pay the bills.

Even though they were 8000 miles away, my family and friends in India have always stood by me and made a difference with their love and prayers. My Mother Jashoda, my father Sundar and my sister Poorvaja have always stood by me in the toughest of times. Their contribution to this work, though unseen, is monumental. Last, but certainly not least, a big hug to everyone belonging to the Indian community here at MSU for always making me feel at home, loved and pampered.

Our research group is forever indebted to the good people of East Lansing who spared their time and energy to appear as test subjects for all our studies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Images in this Thesis are presented in Color

Chapter 1: Warnings and their Effectiveness

Warnings and us

A closer look at our environment reveals that warnings are everywhere. We find them on bottles of alcoholic beverages, on hair dryers and on car seats for children. All around, manufacturers of equipment warn us of potential misuses and hazards. Similarly, manufacturers of drugs warn of possible side effects and adverse events. Warnings are diverse; ranging from those that caution about slippery roads to those that inform about the health risks associated with smoking. There has been an increase in warning research in the recent past (DeJoy and Laughery 1999)

This may be linked to the increasing costs of accidents and healthcare in the US and the escalating importance of warnings in litigation (DeJoy and Laughery 1999). It is also appropriate to mention here that government organizations like the Occupational Safety and Health Association (OSHA), the Environmental Protection Agency (EPA) and the Food and Drug Administration (FDA) have fueled research in warnings in their respective sectors (DeJoy and Laughery 1999). Some see warnings as a means to transfer responsibility of safety to the users (DeJoy and Laughery 1999). That said, warnings are an important means of informing people that hazards exist and actions are needed to counter them. They are an important part of the people's 'right to know'. (Laughery and Hammond 1999)

The sum total of all warnings can be seen as serving a set of three purposes. Firstly, warnings are used to prevent dangerous practices like smoking in gas stations or exposure to sunlight while taking certain drugs. The second purpose is to cause behavioral changes. A warning educating researchers to wear protective gloves is one such example. The third type of warning prompts us to make informed decisions. Required cigarette and alcoholic beverage warnings represent examples of this type.

Warnings can also be classified based on the primary sensory route that they use to warn. Visual warnings can be found on the labels of packages or as visual cues (signboards, strobe lights) in the environment. Warnings may also be auditory; examples of auditory warnings include warning tones, alarms and messages in machines. Olfactory warnings may take the form of chemicals added to odorless gases to help detect leaks. Tactile warnings include textured surfaces on curbs that warn visually impaired people to move away.

Before analyzing warning effectiveness, it is important that the concepts of hazards and their control are understood.

Hazards, Risks and Control

A hazard is defined as *"a set of circumstances that may result in injury, illness or property damage (Laughery and Hammond 1999) "* Risk is the likelihood that a certain adverse event will occur. Following the above two terms, danger is often expressed as the product of a hazard and the likelihood that it will cause damage (Laughery and Hammond 1999). In the context of hazard control, (3) warnings are referred to by some authors as the third line of defense, the first two being better (1) design and (2) guarding. Elaborating, when a different design or practice affords a greater amount of prevention of damage to life and property, it is always preferred over others. Guarding against the hazard is the next line of defense. For instance, heavy machines always have metal plates covering high risk fixtures like cylinder nips and gear trains. Warning about the hazards is the third and final line of defense. (Laughery and Hammond 1999)

After assessing the hazards that may be present in a system (even after effective design and guarding), the likelihood of adverse events is evaluated to determine if the risk is high enough to justify the incorporation of a warning into the system. However, this is done only after thoroughly analyzing the source (the hazard itself), possible channels (visual, auditory, etc) and the receiver (workers, maintenance staff, etc). Previous research also recommends testing of the warning before incorporating it into the system (Laughery and Hammond 1999).

It has been indicated in the literature that properly documented warning and hazard prevention research is scarce because of difficulties in conducting these studies (Wogalter, DeJoy et al. 1999). For instance, it isn't appropriate to expose test subjects to real hazards in their actual setting. In the recent past, studies on warnings and their effectiveness have been fuelled by the development of better study techniques, methods and models (Wogalter, DeJoy et al. 1999).

Warnings are often, if not always, approached with a theoretical framework. The following sections outline a few of the widely used theoretical models used in warning research.

Models and their importance

There are many theoretical frameworks or models that allow researchers to investigate a warning, evaluate its effectiveness or make predictions about the same. These models, though not complete or fool proof in any way, do serve to provide some insight into the effectiveness of warnings before they are deployed. The following paragraphs describe some of these models.

Communication Models: Warnings have always been looked upon as forms of communication. Consequently, the first theoretical models that were used to assess warning effectiveness came from communications theory. In general, modern communication theory has largely evolved from the publication titled, "A Mathematical Theory of Communication" by Shannon, C (Shannon 1948). A communication model analyzes the three main components of the warning process namely, the source, the channel and the receiver (Wogalter, DeJoy et al. 1999). In the case of a package, the source may be the manufacturer of the product. The channel is the warning itself and the receiver is the consumer.

Most forms of the communications model also take into account the effects of noise and feedback. For example, a communications model would be able to take into account extraneous distractions (noise) and also consider alterations in subject's responses based on external cues (feedback). Examples of noise include cluttered visual fields for print warnings and road signs. Examples of feedback include subject responses to auditory beeps/alarms. Certain fields of study that derive from general communication theory like 'persuasive communications' are heavily focused on beliefs and attitudes of consumers who are presented with warnings. An interesting set of studies on warnings on alcoholic beverage bottles and cigarette packs have been carried out using this model (Andrews 1995).

Information Models: Information models, on the other hand, focus on the stages that the information contained in the warnings passes through while being transmitted from the warning to the user. These models take into account facets of information processing like attention, comprehension, attitudes and beliefs, motivation and behavioral change (Laughery and Hammond 1999). Information processing theories often mention that the information passes through several stages in serial fashion and that failure at any stage will lead to failure of the warning process.

A commonly used theory of information processing indicates that the following four steps must be completed for the warning to be effective. (Lehto and Miller 1988; DeJoy 1991; Wogalter and Laughery 1996; Laughery and Hammond 1999; Rogers, Rousseau et al. 1999)

- The information must be **noticed**
- It must be **encoded** into memory. For this to happen the message must be gathered through perceptual systems, generally vision; i.e., the message must be legible
- The encoded message must then be **comprehended**
- It must, finally, move the reader to **action**

C-HIP: Certain researchers stress the importance of consolidating concepts from both information processing and communication theories, creating a common framework that takes into consideration both models. An example of this kind of model is the Consolidated Communication – Human Information Processing (C-HIP) model (Wogalter, DeJoy et al. 1999). The C-HIP combines concepts of both the information processing and the communications models as shown below.

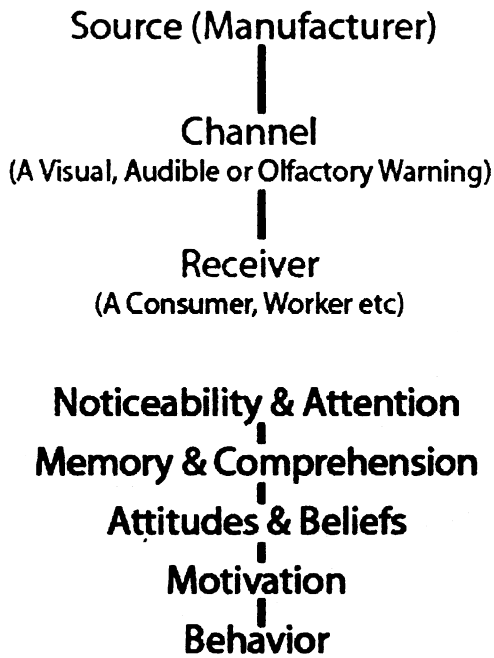


Figure 1 – Consolidated - Human Information Processing (C-HIP) Model

They incorporate information processing into the receiver end of the communications chain. This model is used widely to analyze warnings and organize literature regarding warning effectiveness research. It is noteworthy that entire books have been written on warnings that organize content based on the above C-HIP model.

Warning Effectiveness Research

The following paragraphs will examine some of the variables that are analyzed in warning studies and will also provide an overview about some of the methods that are used to assess the effectiveness of warnings. There are essentially two kinds of warning studies. There are those studies that analyze the warning as it passes through the various stages before it can influence behavior (Type 1) and there are those studies that analyze behavior and behavioral compliance itself (Type 2) (Young and Lovoll 1999).

Type 1: As mentioned before, information models assume that a warning must pass through various stages (attention, sensory routes, comprehension and compliance); these stages are the focus of the first type of study.

The attention phase is analyzed either directly or indirectly. Eye movements, are used as an objective measure of attention and parameters such as fixation time and reaction time have been used to quantify the noticeability of warnings. Indirect measures like recall are also used to measure noticeability. (Fischer, Richards et al. 1989; Krugman, Fox et al. 1999; Crespo, Cabestrero et al. 2007). Other studies based on attention have focused on additions and enhancements to warnings like symbols (Cahill 1976; Swindell 1999), color contrast (Sumner 1932), borders (Ells, Dewar et al. 1980) etc.

After being noticed, the warning must be legible enough to be encoded into the user's memory. This must occur through the different sensory routes. A visual warning must be seen and read; an auditory warning must be heard and understood. In addition to this, the consumer must have the cognitive resources to successfully encode information that is present in a warning. In other words, a warning may be legible, but if the user's mind is already overloaded with other tasks and information, encoding may not occur efficiently.

Comprehension is the next stage in the information processing hierarchy. Studies involving comprehension analyze parameters such as the level of message "readability" (Wong 1999) and the literacy levels of subjects in order to evaluate the likelihood that a person will understand the message (Karwowski, Yates et al. 1991; Davis, Wolf et al. 2006).

Other important concepts that are studied by warnings researchers include attitude, perception and behavior. Research in risk perception aims to determine the attitudes of people as they approach warnings and hazards in general. There have been a variety of studies that involve questionnaires that employ Likert scales with responses involving subjective measures, (Fischhoff 1977; Baber and Wankling 1992). The same questions can also be examined with more objective measures. For example, studies sometimes make use of existing real-world information sets like databases of accident information to quantify risk perception among drivers.

Type 2: As mentioned earlier, the ultimate goal of a warning is to bring about a behavioral change in its intended audience.

"If a warning is effective at the behavioral stage, the warning is probably adequate at the earlier stages. Indeed, behavioral data are so important that if only one measure of warning effectiveness can be obtained, a compliance test is the best one to do. (Wogalter and Dingus 1999)"

These studies may be divided into those that track actual behavioral compliance and those that analyze behavioral intentions alone. There are several problems that prevent direct measuring of behavioral compliance. Some of them are mentioned below

- It is not ethical to expose participants to real hazards. Such studies are also difficult to set up when the real hazards are unpredictable (natural disasters).
- They are often time consuming and difficult to manage (Wogalter and Dingus 1999).

- Studies that examine direct measures of behavioral compliance require very careful planning. Incidental exposure may be critical (tell the participant that the study concerns something else).
- Furthermore, there must be careful screening of participants. Researchers must make sure that studies are conducted on subjects that are likely to benefit from the incorporation of an effective warning, i.e. the test population must be carefully selected and monitored.
- Small nuances in the experimental design will affect the outcome of the study and, therefore careful planning of the experimental design is required (Wogalter and Dingus 1999).

During this literature survey, four broad categories of behavioral compliance experiments were identified. They are described below with an example each.

- **Self reports** of compliance generated through questionnaires and interviews are sometimes conducted to examine the effectiveness of warnings. The difference between these and studies that measure behavioral intentions, is that these questionnaires are administered "*ex post facto*" (Wogalter and Dingus 1999) , i.e. the questionnaire is given to the subjects after the actual response has taken place.
- Studies in warning effectiveness sometimes involve actual **Observations of behavioral compliance** Examples include studies in chemical laboratories where warning effectiveness and compliance behaviors of researchers has been analyzed based on actual observations of their activities (Wogalter, Godfrey et al. 1987).
- Other studies that measure behavioral compliance involve innovative methods of **Physical Tracing**. Studies of this type involve some kind of measure that is used

to detect whether the participant paid attention to the warning and complied with its message. One such study assessed the effectiveness of a warning advising usage of gloves while spraying cleaning solution. The gloves were inspected to see if they had been worn while spraying by observing physical damage to the fingertips of the glove from the spray nozzle (Hunn and Dingus 1992).

- **Epidemological compliance studies** concentrate on warning effectiveness across large populations. Studies that have evaluated the effectiveness of driver training programs in preventing accidents across populations are examples of this type (Edwards and Ellis 1976).

Studies on behavioral intentions, as opposed to behavioral compliance, are more widespread. Research indicates a strong link between intentions and compliance, i.e. there is a link between what people intend to do and what they actually do (Ajzen and Fishbein 1977). Behavioral intentions are often assessed using Likert scales and questionnaire responses. The above summary of methods and experiments is by no means exhaustive, but it does put things in perspective.

Warnings and the Law

As mentioned before, it is likely that recent warning research is motivated by widespread lawsuits and increasing settlements related to warnings, their absence or ineffectiveness. The concept of U.S. product liability says that a manufacturer who designs and sells a product must warn users of potential problems that arise during its foreseeable use and instruct them on its safe usage (Madden 1999). The duty of a manufacturer to warn about a product is explained under two principles: Negligence and Strict Liability.

Negligence: The principle of negligence says that when there is any foreseeable harm that may occur during the use of a product, the failure to warn is termed as negligence and is grounds for a lawsuit (Madden 1999).

Strict Liability: The principle of strict liability in tort states that a manufacturer has the duty to warn when a normal person cannot perceive the dangers associated with a product. This is to say that when the dangers of a product go beyond what would normally be foreseen by a user, the manufacturer has strict liability in a tort litigation. There also exists a rule that states that a manufacturer need not warn at all when the risks and hazards are obvious. For instance a knife need not contain a warning about a sharp edge (Madden 1999).

A full analysis and review of the various laws and practices related to warnings is beyond the scope of this chapter. A limited list of exemplar issues in lawsuits involving warnings follow.

- Obviousness of risks and hazards play an important role in litigation. Case studies indicate that lawsuits initiated by users of BB guns, darts and slingshots have been rejected because the dangers of these products were obvious (Madden 1999).
- An automobile manufacturer was once given a favourable judgement when a person accused them of not providing a warning. The person had tried to prime a carburetor with gasoline poured from a quart jar (Madden 1999). The manufacturer was able to establish that the act was careless in itself and the user would not have heeded the warning anyway.
- There are cases where the charges are transferred to a third party. For instance, in hearings that involve prescription drugs, it is common to transfer the liability from

the manufacturer to the physician who serves as a "learned intermediary." (Madden 1999).

- Several court cases have shown that a manufacturer must be able to foresee possible alternate uses of the product and warn against the same (United States Court of Appeals 1974). One such example involved a housewife who stood on a chair when the chair's backrest failed. The manufacturer was penalized for the same (Madden 1999).
- Manufacturers are expected to warn (FDA 2004) about ingredients in their products that cause allergy in substantial proportions of populations, e.g. Peanuts.
- There have been cases in which "For professional use only" warnings were not heeded by household users (Madden 1999). There was once a case where a professional hair care preparation was not used properly by a household user and this resulted in injury. The verdict, however, was in favor of the manufacturer. The hazards of the product would have been obvious to professional beauticians.
- There have been disputes about warning effectiveness with reference to the design and layout of the warning itself. A manufacturer of furniture polish was found guilty for printing the warning on its package in the same color and typeface as the rest of the text. This resulted in the subsequent death of an infant who injested the product. The warning in this case, was not considered adequately noticeable and did not effectively convey the gravity of the hazard (Madden 1999).

Warnings in Packaging

Several authors mention that communication is one of the primary functions of any package, along with protection and containment (Paine and Institute of Packaging.

1963; Soroka and Institute of Packaging Professionals. 1995). Paine (1963) mentions in his writing,

"It is obvious that retail packages must communicate, for not only do they have to identify the contents, but in many instances they have to sell them as well (Paine and Institute of Packaging. 1963)"

With risks comes the need and requirement to warn. Food packages may contain warnings about allergy causing ingredients, storage conditions, contamination indicators and tamper evidence. Medical packaging often contains warnings about side effects, risks and possible drug interactions. Warnings on packages may take many different forms, ranging from simple labels and print warnings to information inserts that may be several hundred words long. In the context of logistics and distribution, packages contain warnings about storage and handling.

Chapter 2: Prescription drug warning labels (PWL's)

Patients and medical information:

It is estimated that around 52% of all Americans take at least one prescription drug every week. This figure rises to 82% when we add non-prescription drugs into the equation. Older age groups have a higher per capita consumption of prescription drugs. The Slone survey of drug usage in the American population for 2006 says that 17-19% of the people in the age group of 65+ take at least ten prescription drugs in a week (Kaufman, Kelly et al. 2006). As a result, drug regimens in older populations tend to be more complex than their younger counterparts. Thus, there is great concern about the management of these therapy regimens in older populations (Gryfebecker, Segal et al. 1989).

Further, as healthcare shifts from the inpatient setting to the outpatient setting, furnishing timely information and achieving patient compliance becomes more difficult (Davis, Wolf et al. 2006). Studies have shown that a significant proportion of patients do not comply with instructions that are given about prescription drugs (Rehder, Mccoy et al. 1980; Boyd, Boyd et al. 2006). It has been estimated that nearly 1.5 million medication errors occur annually in the USA, and that a majority of them are in the outpatient setting, where it is up to the patient to use the information provided (Bates 2007) to him or her about the drug. It is also indicated that 33% of these errors cite packaging and labeling as the primary cause (Berman 2004). It can also be gleaned from the literature that medication errors, in general, generate costs of almost 13 million dollars annually (Aspden and Institute of Medicine (U.S.). Committee on Identifying and Preventing Medication Errors. 2007).

Auxiliary labels and PWL's:

Pharmacists have often employed verbal counseling to inform patients about drugs, their side effects and dosage instructions. There are studies to suggest that this type of verbal instruction increases patient compliance (Ascione and Shimp 1984; Johnston, Clarke et al. 1986). However, there are disadvantages to verbal counseling. Issues with memory and patient comprehension can result in error, even when verbal instructions are provided (Gryfebecker, Segal et al. 1989). Thus, pharmacists depend on printed information to supplement verbal counseling. These may take the forms of Patient Information Leaflets (PIL's) and patient packaging inserts (PPI's) as well as 'auxiliary' labels that are affixed on prescription vials. Sometimes these auxiliary labels exist as the small colorful, prescription warning labels (PWL's).

In packaging as a whole, labels assume great importance from an information perspective as they always stay with the packages and provide timely information about the product (Miller 1978). Yet, studies have shown that the white pharmacy labels (see Figure 2) attached to the vials do not provide adequate information about the drug when used alone. This is one reason why the auxiliary label is important (Mazzullo, Lasagna et al. 1974). Authors have mentioned that the prescription warning labels (PWLs), aim to convey information about appropriate medicine use, potential drug interactions and side effects in the simplest way at the least possible cost (Brown, Solovitz et al. 1988).

Prescription drug warning labels (PWL's) contain warning messages (see Figure 2) such as, "Do not consume alcohol while taking this medication" or information about routes of administration, viz. "For external use only." At present, the required information for prescription drug warning labels and the presentation of said labels is not standardized in the USA, and they are not regulated by the FDA (Ault 2007). This is despite the fact that, in many cases, these labels contain information that is critical to the safe and effective use of the products contained within. Furthermore, studies have revealed that many consumers do not heed these warnings (Franklin 2005; Davis, Wolf et al. 2006).

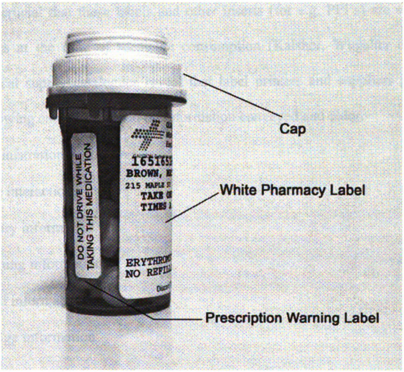


Figure 2- A photograph of a vial used in the study. A common 10 dram prescription vial of the 1 Clic type with two types of labels; PWL's and White Pharmacy Labels.

Figure 3 – Examples of actual PWL's used in local pharmacies

It is crucial that these labels and other inserts (for e.g. PPI's) are effective and warn patients at the time of medicine consumption (Kalsher, Wogalter et al. 1996). Pharmaceutical supply catalogues reveal that label printers and suppliers group labels into the following categories based on information conveyed and color.

- Administration Route
- Drug interactions
- Dietary information
- Warning information
- Refill information
- Dosage information
- Storage instructions

There have been several studies on these crucial warning labels and their effectiveness. The following section provides an overview of some of them.

Studies on PWL effectiveness:

Over the years, there have been studies that have looked at prescription warning labels from several aspects ranging from reading difficulty (Davis, Wolf et al. 2006) to symbol presence or absence (Wiederholt, Kotzan et al. 1983).

A study conducted as early as 1983 by Wiederholt et al (Wiederholt, Kotzan et al. 1983) looked at the effectiveness of these labels using a tachistoscope. A tachistoscope is a mechanical device that was used to display images in studies that evaluated attentive behaviors. In this study, the tachistoscope was used to display images of prescription vials with PWLs pasted on them. The experiment measured recall through multiple-choice questionnaires. The study looked at the relative effectiveness of color and black/white labels and the presence and absence of a graphic symbol on the recall of warning information. An interesting experimental condition in this study was the incorporation of the warning text on the pharmacy label itself, i.e. the non-usage of a warning label. Researchers concluded that there were significant increases in recall when the PWL was used, but also elaborated that the study did not conclusively evaluate the effects of color and symbols.

In the same year Wilson and Hogan conducted an analysis of PWL text and concluded that 11.2 percent of the 110 words analyzed in this study were above the 8^{th} grade reading level. They also suggested the use of a 'word familiarity index' to test the readability of these labels as opposed to using the traditional readability tests. (Wilson and Hogan 1983)

Ferguson et al (1987) conducted a study on warning labels that looked at the various levels of severity in warning label text. Subjects were questioned on their perception of different levels of strength/severity of warning label text. Authors conclude that patients could be made to comply with therapeutic regimens by increasing the severity of warnings. Subjects also perceived the medium severity warning to be the one most effective (Ferguson, Discenza et al. 1987).

Other early studies by Brown et al. (1988), looked at the effect of these labels on the memory of the patient for drug related information. It was concluded that PWLs were an effective means of educating patients about drug related hazards and were very effective on a short term basis. They also discussed the phenomenon of generalization of medical information that occurs with patients, i.e. a patient being told to avoid aged cheeses may begin to avoid all kinds of cheese (Brown, Solovitz et al. 1988).

It has been mentioned before in this chapter that medication use is higher among the elderly. Early studies in 1989 by Gryfe-Becker et al, investigated the knowledge that elderly patients would gain from auxiliary labels. The study specifically used auxiliary labels as supplements to verbal counseling and compared the rates of drug information recall among patients who received only verbal counseling while being given bags with drug vials as opposed to those who were also provided with auxiliary labels on the vials. The interviews to measure recall information were conducted 10-15 days after the medication was picked up by the patients from the pharmacies. It was concluded that those subjects that were given auxiliary labels were able to recall drug information at the higher rate than those who were not (Gryfebecker, Segal et al. 1989).

A study published in the Journal of Pharmacy Technology in 1991 focused on pharmacists and their patterns of PWL use in the state of North Carolina. The results of this study showed high levels of non-usage of PWL's among several pharmacies for many drugs. The usage of PWL's differed between pharmacists employed in chain stores, pharmacists using label software, and those who worked in independent stores (Dombrowski, Schuna et al. 1991).

A recent study on designs and pictorials in PWL's analyzed tagged and fold-out labels as alternatives to the existing designs. For instance, fold out labels required that subjects unfold them from the side of the container and then down. This label had a surface area that was 40 % greater than that of the conventional label. Subjects were asked to rate designs after being exposed to them. The study also looked at the presence of pictorials on these labels and their effectiveness. The study reported a higher preference for the new designs and labels with pictorials among all age groups (Kalsher, Wogalter et al. 1996).

A widely cited and quoted study on the reading difficulty of PWL's used the REALM literacy test to determine health literacy levels among subjects and LEXILE scores to measure the 'reading level' of the label text. The study attempted to correlate readability levels as measured using the LEXILE text and patient understanding of the warnings themselves (Davis, Wolf et al. 2006). It was concluded in this study that those subjects who were reading at a low literacy level were less likely to understand and interpret the meanings of the PWL text correctly. Around one-third of the total number of subjects tested were below a sixth-grade reading level.

In subsequent sections, other important aspects of this study such as eye tracking, aging and health information shall be analyzed.

Chapter 3: Aging and the Warning Process

America's seniors and Medical Information:

In a 2006 report written for the National Institutes of a Health (NIH), it is mentioned that the population of the United States of America over 65 years of age is expected to double in the next 25 years. It is projected that 20% of the US population will be aged 65 years and over by 2030 as compared to 12 % in 2003 (NIH 2006). A deeper exploration of related literature suggests that there has been a concurrent increase in the number of citizens seeking medical care. A 2008 article published in USA Today shows that the number of medical visits increased almost 26% between 1996 and 2006. The article goes on to mention that the aging population may contribute to increased medical visits (Hellmich 2008). Data released by the US Census bureau suggests that the baby boomer generation will further tax American healthcare as they age. Quoting from the US Census bureau source,

"*Just as this generation had an impact on the educational system and the labor market, this large cohort will strain services and programs required by an elderly population.....Between 1990 and 2020, the population age 65 to 74 would grow 74 percent under middle series projections, while the population under age 65 would increase only 24 percent. (Bureau 1997)*"

Other sources indicate that the proportion of seniors living alone is also increasing (Report 2003). As stressed previously in this document, the approach to healthcare delivery is shifting from the use of inpatient to outpatient settings. All these factors point towards the increased need for efficient design of health information and its delivery.

Warnings and Aging – Physiological and Cognitive changes:

Dejoy asserts that there has not been significant warnings-related research in which age is studied as a direct variable of interest (DeJoy 1999). This section will attempt to summarize some of the physiological and cognitive factors that vary with aging and how they affect the process of warning. This section draws information from and owes its structure to the book chapter, "Maximizing the effectiveness of the warning process: Understanding the variables that interact with age" by Gabriel. K. Rousseau and Wendy. A. Rogers (Rogers, Rousseau et al. 1999).

Rousseau and Rogers mention that there are 'person' variables and 'warning' variables that play a part in warning effectiveness. 'Person' variables include those that depend on the receiver, age being one of them. Age in turn affects other person variables such as vision, cognitive capability and memory. Warning variables include characteristics of the warning such as color, typeface and type size (Rogers, Rousseau et al. 1999).

Several studies reveal that there are important 'person' variables that change with age. Some of them are mentioned in Table 1. A short overview of each factor in relation to warnings and age is mentioned in the following sections.

Physiological Variables

Vision: Studies have shown that several facets of vision decline with age. Existing literature has documented declines in visual acuity, contrast sensitivity, acuity for peripheral targets and color discrimination (Kline and Scialfa 1997). This reduction in visual capability suggests that older individuals will have trouble noticing and comprehending warnings. The basic visual acuity of subjects also declines sharply with age (Bix 1998). Literature reviews of vision in senior citizens assert the need for increased contrast in warnings in products and systems that are intended for older audiences (Charness and Bosman 1992). Studies looking at older adults have constantly stressed the need for senior-friendly features in the warnings, such as matte surfaces to reduce glare and use of fonts that are over 14-points in size (Backinger and Kingsley 1993).

There is some literature that attempts to quantify the changes in color perception as human beings age. A guidance document released by the Department of Health and Human Services titled "Recommendations for developing user instruction manuals for medical devices for home health care", mentions that the color blue is very difficult for older adults to see and should be avoided (Backinger and Kingsley 1993).

Gender: It is acknowledged that gender does play an important role in influencing an individual's willingness to look for warnings. Studies by different researchers have established that females have a greater tendency to look for warnings than males when a product is perceived to be hazardous (Godfrey, Allender et al. 1983; Goldhaber and Deturck 1989). Rosseau and Rogers mention that there is yet no literature available to predict how these gender differences behave with age (Rogers, Rousseau et al. 1999). However, interesting warning related phenomena like the 'boomerang effect' explained later in this chapter have been found to depend on a combination of age and sex.

Table 1 - Person Variables that effect Warning effectiveness

<table>
<tr><th colspan="12">Physiological Variables</th></tr>
<tr><td colspan="6">Vision</td><td colspan="6">Gender</td></tr>
<tr><th colspan="12">Comprehension Capability</th></tr>
<tr><td colspan="6">Reading Comprehension</td><td colspan="6">Symbol Comprehension</td></tr>
<tr><th colspan="12">Cognitive Factors</th></tr>
<tr><td colspan="3">Familiarity</td><td colspan="3">Hazard Perception</td><td colspan="3">Costs/Benefits of Compliance</td><td colspan="3">Risk Taking Style</td></tr>
<tr><td colspan="4">Attention</td><td colspan="4">Working Memory</td><td colspan="4">Prospective Memory</td></tr>
</table>

Comprehension

Reading Comprehension: Studies comparing college-going students to middle-aged adults have assessed the "willingness to read warnings" at different ages (Silver and Wogalter 1990). These studies have reported that the general willingness to read warnings as reported by subjects does not vary greatly in between these two age groups, but there are several significant declines in the ability of subjects to process information gleaned from text. There are studies on general prose comprehension that report that older adults are unable to make technical inferences and tie in information from different parts of the same text (Kemper 1992).

There are also important age related declines in functional health literacy. (Williams, Parker et al. 1995). Tests have shown that between 48 to 80 % of adults over

60 years of age have inadequate functional health literacy (Williams, Parker et al. 1995). In a study conducted in 1993 by Hill, findings indicated that 18% of a sample of 83 community dwelling, educated, elderly adults "*were reading below a twelfth grade level on the reading subtest of the Adult Basic Learning Examination (Level 3)*" (Hill 1993). Other sources indicate that age related declines in reading comprehension are not very significant for highly educated older adults (Meyer, Young et al. 1989; Harris, Rogers et al. 1998).

Symbol Comprehension: Symbols, graphical and iconic representations of concepts, actions or instructions are important in warning design and effectiveness. Some authors choose to distinguish between symbol (icons) and pictorial (actual images) comprehension, but most research studies do not distinguish between the two (DeJoy 1991). As mentioned before, symbols are very important in conveying warning information when consumers are from varied linguistic backgrounds. In general, studies have shown that the use of a symbol-text combination in a warning always yields better compliance behavior among consumers than the usage of a symbol or textual warning alone (Otsubo 1988; Janynes and Boles 1990; Young and Wogalter 1990). In terms of the differences in symbol comprehension across different age groups, there is some evidence to suggest that older adults do not recognize and understand symbols as well as younger adults, but this is still very vaguely understood (Rogers, Rousseau et al. 1999) .

Cognitive Factors

Familiarity: Familiarity is discussed by Rosseau and Rogers under three subtopics (Rogers, Rousseau et al. 1999)

- General Experience

- Schemata
- Technical Knowledge

General Experience, as the term suggests, deals in past experiences with the product or system. It is generally understood that when people have used a product or system for a long period of time with no adverse events, they tend to overlook warning information (Godfrey and Laughery 1984). Studies among different age groups (under 30, 30-50 and 50+) have revealed that most individuals wouldn't look for warning information if they had used the product before without harm. Subjects have sometimes reported that they were more likely to look for and comply with warning information in electrical products than non-electrical products (Wright, Creighton et al. 1982).

Schemata are general knowledge structures that individuals rely on to explain and predict events in life. One particular study analyzed the schemata of different age groups relating to medical information by instructing subjects to arrange random blocks of medical information *"into the [perceived] optimal order for effective instructions"*. The optimum schema was determined for different age groups and these ordered instruction sets were used with a different set of subjects to assess recall. The results of the study mention that the ordered information sets resulted in increased recall as opposed to original information sets for all age groups (Morrow, Leirer et al. 1996).

Studies have shown that warning information is better recalled when presented first in a set of information (primacy) and last (recency). Primacy (presenting information in the beginning) and Recency (presenting information at the end) of information has a tendency to promote recall, even though it seems to go against ordering schemata as analyzed above (Wogalter, Fontenelle et al. 1985).

Technical Knowledge of drugs and medical information, in general, is low among older adults. *"Expecting older adults to comprehend detailed technical information presented on warning labels may be unreasonable (Rogers, Rousseau et al. 1999)".*

Hazard Perception: A survey of existing warnings related literature reveals that consumer attitudes and beliefs towards warnings vary with age (DeJoy 1991). There have been studies that have looked at the attitudes of consumers from different age groups (college, high school and middle school) to "signal words" like: danger, and caution. Study results suggest that the youngest consumers generally exhibited a higher perception of danger than older consumers when these signal words are present (Silver and Wogalter 1990). These findings are consistent with others that relate signal word and age to hazard ratings, with younger consumers assigning higher hazard ratings than their older counterparts (Wogalter, Jarrard et al. 1994). A study by Wogalter et al (1994) asked subjects (high school students, college students and shopping mall visitors aged 21-80) to rank warning labels found in actual consumer products in terms of perceived hazard. The results of the study showed that the rank ordering was consistent among different age groups. Additionally, it has been indicated that when consumers perceive a hazard to be very high, they are more likely to notice, encode, comprehend and comply with the associated warnings (DeJoy 1991).

Cost/Benefits of Compliance: There are always certain 'costs' to compliance (e.g. wearing safety glasses, locating them, wiping them clean and so on). It would appear that older adults and seniors would not comply with warnings when the cost of compliance is too high. In other words, seniors may not have the physical and cognitive resources needed to comply with a warning that has a high and complex compliance cost (Rogers 1983).

Risk-Taking Style: Case studies have investigated the likelihood of younger adults to voluntarily assume risk by subverting warning information (DeJoy 1991). This is sometimes referred to as the 'boomerang effect', named after the Australian hunting implement that returns to the user when thrown. A study by Goldhaber and deTurck compared the likelihood that middle school and high school students would comply with a 'no diving' sign near a swimming pool. In this study it was seen that males from high school were more likely to indulge in the prohibited behavior (boomerang effect). However, another finding from this study was that high school female students were more likely to comply with the warning sign, thus indicating an effect of gender on the boomerang effect (Goldhaber and Deturck 1989). Other studies comparing middle aged to young adults indicate that middle aged adults (40+) were more likely to comply with warnings in general (DeJoy 1999). So far, studies have shown that older adults very rarely take risks voluntarily.

Attention: Even though there aren't many studies that have looked at the differences of attention in relation to warnings and aging, there is literature to suggest that attention, by itself does decline with age. There are studies to show that both divided attention (ability to multi-task) and selective attention (ability to gain cues from the environment) show significant reduction with age (Rogers, Rousseau et al. 1999).

Working Memory: It has been mentioned before that working memory, or short term memory, declines with age (Salthouse 1991), and that older adults have trouble comprehending warning label text when it overloads the working memory. Rousseau and

Rogers (1999) mention that working memory overloads may cause problems in all four stages of the information processing model. Warnings may not be noticed, may not be encoded properly when noticed, may be difficult to comprehend and so on. One particular project by Morrell et al (1989) studied complex medical regimens and their relation to declines in the comprehension of medical information (Morrell, Park et al. 1989).

When warnings and health information overload the subject's working memory, there is always a risk of the subject not being able to comprehend the information on the warning. An example of warning text that overloads working memory is as follows,

"Consult a doctor if symptoms persist or get worse, if new ones occur, or if sleeplessness persist continuously for more than two weeks because these may be symptoms of a underlying medical illness (Rogers, Rousseau et al. 1999)"

Analyzing this piece of text, it is obvious that the multilayer (if...or if...because) structure of this sentence leads to an overload on the working memory of the subject and causes problems in information processing.

Another study on comprehension and understanding by Zuccollo and Liddell in 1985 concluded (Zuccollo and Liddell 1985) that

"60% of elderly patients (mean of about 79 [years of age]*) had trouble reading their medication labels and only 23% had a clear understanding of all the instructions (Rogers, Rousseau et al. 1999)"*

Prospective Memory: "Prospective memory means remembering to perform actions in the future (Rogers, Rousseau et al. 1999)". Certain warning-related tasks may be event based or time based, requiring delayed action for proper compliance. Research exists to show that older adults have a problem with time-based tasks. However, Rousseau and Rogers mention that there is 'remarkably little research' on prospective memory and its relation to warnings in general.

Chapter 4: The Human Eye and Eye Tracking

The Human Eye:

An understanding of the basic structure and working of the human eye is necessary to understand how eye tracking works, and how the corneal reflection technique uses the structure of the eye to its end.

The eyeball is surrounded by a tough and opaque membrane known as the sclera. There is however, a small disc shaped transparent region in the front called the cornea. It is through the cornea that light enters the human visual system. Underneath the membranous structure of the sclera is a vascular layer that has a dense network of vessels that supply the eye with blood and other fluids.

The sclera does not stop with the underlying vascular membrane. The vascular membrane surrounds the innermost part of the eye known as the retina. This is the layer that serves as the eye's primary visual sensor. The retina is rich in light sensitive cells. These cells are divided into rods (light sensitive) and cones (color sensitive). The retina does not have an equal distribution of these cells throughout.

The point near the entry of the optic nerve into the eyeball is known as the *macula lutea* and within this region is the *fovea centralis*. The fovea centralis is associated with the point 'of clearest vision' and is composed mainly of cone cells. The fovea is about 0.4 mm in diameter (Yarbus 1967).

The retina itself is a highly complex and layered structure with multiple layers of nervous, vascular and optic material. The nervous cells in the inner layers of the retina all lead to the optic nerve that functions as the carrier of impulses between the brain and the eye. This nerve is about 5 cm long and around 4 sq. mm in cross section (Yarbus 1967).

Before light can reach the optic nerve via the retina, it must first be introduced through the cornea. The cornea transmits light into a lens arrangement whose focal length is controlled by a ciliary body. A paired set of muscles known as the iris controls the aperture of the lens. This aperture of the lens is known as the pupil and diametrical changes to the pupil are involuntary. The diameter of the entire structure of the eyeball is, on an average, 24 mm measured in all meridians (imaginary lines drawn through the center of the eye). As shown in the diagram, the rectus muscles control the movement of the eyeball and enable focusing on objects in a wide field of view (Yarbus 1967).

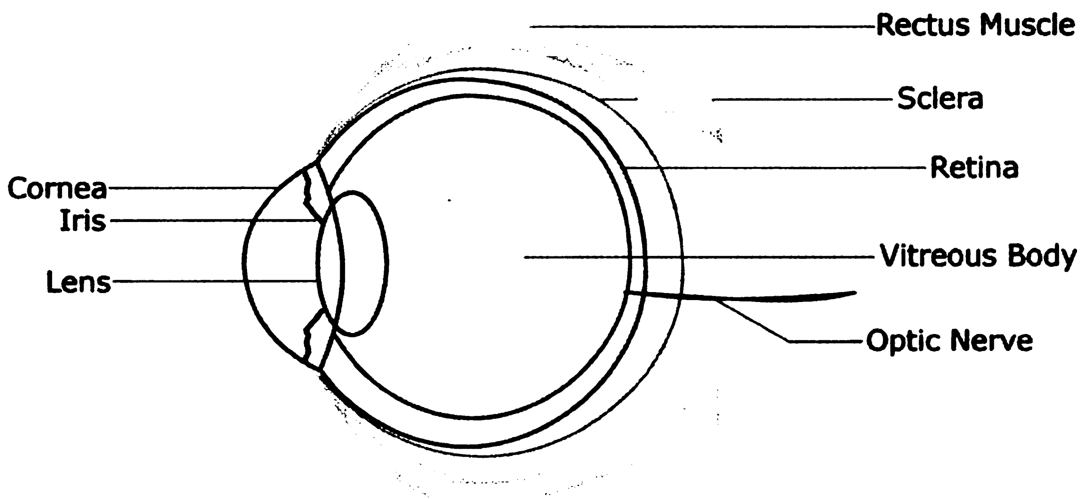


Figure 4 – A schematic diagram of the structure of the Human Eye

Eye Movements:

From an anatomical and psychological perspective, eye movements do more than just shift the point of focus of the visual system. Research has shown that an important requirement for vision is the presence of a non-stationary retinal image, i.e. an image on the retina that is not constant (Yarbus 1967). Subsequently, there is always some movement of the eye needed in order to achieve the shifting image. If by some artificial means, a stationary retinal image is created, the eye loses its power of resolution and ceases to see (Yarbus 1967). Voluntary and involuntary movements of the eye serve various functions in the human visual system. The two most basic eye movements are fixations and saccades. Along with fixations and saccades, drift and tremor are the two other forms of eye movement. All four movement types are breifly introduced below.

Saccades: Saccades are rapid sweeping movements used to shift the point of visual focus. Saccades take very little time to execute, lasting between 10milliseconds and 100milliseconds. During this movement the user is unable to see. Some special saccadic movements are described as 'smooth pursuit' movements where the eye tries to follow a moving a target and the saccade seems to match the velocity of the moving object. Saccades can be both voluntary and involuntary. Saccades usually end in fixations (Duchowski 2003).

Fixations: Fixations are stabilizing movements that enable to eye to focus on a particular point in the visual field.

Drifts: It was mentioned in an earlier paragraph that there should be no stagnation of the image that is formed on the retina for normal vision. Building on this concept, It has been suggested that the eye never has a constant point of fixation (Yarbus 1967). Instead, there is a fixation field that is created due to drifting movements of the eyes that occur during fixations. These drifts serve to keep the retinal image non-stationary during fixation and maintain the resolving power of the eye at an optimum level.

Tremors: Tremors are barely perceptible movements of the eyes that require very highly sensitive eye tracking methods to resolve and observe. Tremors are characterized by extremely quick and small movements of the eye and they occur because of continuous internal nervous activity (Sheahan, Coakley et al. 1993).

Out of the above movements described, fixations and saccades are the most important to visual gaze. Fixations are used determine points of attention and saccades are used to study the gaze trail.

Methods to study eye movements:

Over the years, there have been many instruments and procedures used to study eye movements. Some of the earlier methods were very intrusive and elaborate setups, but most methods used for eye tracking today are non-intrusive and rely on arrangements that have no direct contact with the eye itself. A brief description of some of these methods follows:

After Images and Observational Studies: Early methods used to study eye movements were those that used retinal 'after-images'. Generation of these images on the surface of the retina involved the use of blinding flashes of light. Once the after-image was created, a trained subject could observe this image in his or her field of vision. By studying the movement of this after image, eye movements could be monitored as these images were strictly stationary with respect to the retina. Methods to study eye movements based on simple visual observations were also practiced at this time. After images are no longer used in eye movement research, and they have been mentioned here merely for their historical significance (Duchowski 2003).

Mechanical Methods: There are several methods to record eye movements by mechanical means. The earliest accurate eye tracking methods involved the use of suction devices attached to pointers on the surface of the eyeball. As the eye moved, the pointer would move and trace paths. There were some later studies that used a camshaft mechanism on the cornea. When the cornea moved, the cam would be displaced and it would move a pointer that traced a line on a sheet of paper. The advantages of these methods are extreme accuracy and fine detail. Even barely noticeable eye movements like short saccades and tremor were detected with ease. However, these methods are very intrusive and cause a great deal discomfort to subjects. Furthermore, the preparation of the equipment required for testing was tedious and elaborate (Duchowski 2003).

Electric Methods: A very small electric voltage is always present between the outer and inner sections of the retina and as well as between the cornea and the sclera. As the eye moves, this potential difference varies. Electro-oculography is a technique that employs this phenomenon to track eye movements. Electrodes similar to those used for ECG measurements are placed on specific points on the forehead and the potential created is recorded (Duchowski 2003).

Optical Methods: The majority of inexpensive eye tracking methods in use today rely on optical methods.

Many optical methods, such as the one used in this study, employ the phenomenon of corneal reflection (CR). The cornea of the eye reflects a portion of the light that falls on it, as does any transparent curved surface. The 'bright spot' that is created on the cornea is slightly displaced from the center of the eye ball. The relative positions of this CR spot and the eyeball are used in many techniques to study eye movements. CR based eye trackers are not intrusive, are inexpensive and are easier to setup and use than mechanical methods. However, one major disadvantage of using the CR to perform eye tracking is that the CR also changes position with the head. Stillness of the subject's head is an important pre-requisite in setups that use the CR to track eye movements (Duchowski 2003). This might be difficult for people with neurological or musculoskeletal conditions that make it difficult to remain still. Other optical systems include those that use the Purkinje images that are formed on the eye ball.

Eye Tracking – Applications and Spheres of interest.

Eye tracking has been employed in several fields to study attention, natural behavior and design. Duchowski, in his book on eye tracking methodologies and practice, divides the applications of eye tracking into interactive and diagnostic applications (Duchowski 2003); diagnostic applications are those that help researchers gain some insight into the attentive behavior of the subject. This section attempts to summarize some of the major diagnostic applications of eye tracking and provides references for further reading.

Neuroscience and Psychology: Eye tracking has found vast applications in these fields. Studies in attentional neuroscience have researched the relation of eye movements to the other sensory inputs such as hearing and smell (Asaad, Rainer et al. 2000). Other studies involve analyzing neural activity in the brain in relation to eye movements like saccades and drifts (Gamlin and Tweig 1996; Snodderly, Kagan et al. 2001).

There is also a widespread body of research on reading and visual cognition. Eye movements have been used extensively to study how human beings read lines of text (Rayner 1992). Studies on concepts like scene perception have also increased our understanding of how people scan over scenes and glean information. Studies on eye movements contribute to the knowledge of Schemata (sequences of steps) that are used by people when they fixate over a scene filled with visual information. There are also eye tracking applications in fine arts and visual media. A famous large-scale study at the National Gallery of London used eye tracking to study subjects as they fixated over digitized representations of paintings (Wooding 2002). Similar studies have been carried out on other media such as film (d'Yedwalle, Desmet et al. 1998).

Other diagnostic studies in psychology using eye tracking include studies on illusions, problem solving, visual search and even brain damage and its effects (Duchowski 2003).

Industrial Engineering and Human Factors: *"Eye tracking offers a unique measure of human attentional behavior. This is particularly important in evaluating present and future environments in which human do and will work (Duchowski 2003)"*

Eye tracking has been used many times in evaluating human performance in modes of transport such as aviation and driving. In aviation, eye tracking is used in studies ranging from flight simulation to evaluation of graphical displays. Control panels of aircraft and their ease of use are also studied using eye tracking (Anders 2001). Loss of visual attention is attributed as the cause of a majority of road accidents (Chapman and Underwood 1998). Studies on visual attention of drivers using eye tracking are, thus, very important.

In industries that rely on visual inspection like meat, poultry, electronics and quality control, eye tracking has been used to study and determine optimum methods for the same (Duchowski 2003).

Marketing and Advertising: There are widespread applications of eye tracking in the field of persuasive communications. A study by Lohse in 1997 applied eye tracking effectively to the field of copy testing and analyzed consumer eye gaze while fixating on print ads in the Yellow Pages ™ (Lohse 1997). Eye tracking has also been used to study print advertising and investigate phenomena like "wear-out" associated with repeated exposure to print advertisements (Rosbergen, Wedel et al. 1990). A widely cited study on warning labels in cigarette advertising used eye tracking to assess adolescent attitudes towards them (Krugman, Fox et al. 1994).

The above paragraphs provide only a brief overview of the possible diagnostic applications of eye tracking. These serve to place the method in perspective. The reader is directed to the references section to gain further insight into the studies mentioned above.

Packaging: It is obvious that eye tracking can be applied in a variety of scenarios important to packaging designers and engineers. In a retail environment, it is important to understand how consumers 'view' supermarket aisles and how they scan packages before they make purchases. Indeed, the potential of eye tracking as an important tool for research in retailing has prompted major firms such as Unilever, Walmart, Kimberly-Clark, Target and General Mills to invest in and conduct studies using eye tracking systems. It is noteworthy that some of these setups have received coverage on popular news channels (Ahlbaum 2008).

Closing Statements:

The information contained on the labels of pharmaceuticals, are critical to their safe and effective use. This study attempts to use eye tracking to quantify the noticeability and legibility of PWL's. Other ongoing studies at the School of Packaging, Michigan State University, USA attempt to study overt anti-counterfeiting features and change blindness.

Chapter 5: Print Legibility

Introduction:

Several authors have stressed that communication is one of the primary functions of packaging, (Paine and Institute of Packaging. 1963; Soroka and Institute of Packaging Professionals. 1995) and that warnings are an important form of communication (Wogalter, DeJoy et al. 1999). It is known that very often, packages have warnings that are critical to the well being of the consumers. The four stage information processing model that was used to explain warning effectiveness in an earlier chapter stresses that warnings must be legible in order to be encoded into the memory of the receiver. In visual warnings, this encoding can occur only if the printed message is legible (Rogers, Rousseau et al. 1999) .

A quotation from McLean, as quoted by Pietrowski (1993), clearly indicates the complexities associated with legibility (Pietrowski 1993).

"Legibility is a dangerous – and interesting – word. It is dangerous because it is so often used as if it had a definitive or absolute meaning, which it does not. It is a personal word neither scientific nor precise. If you say 'that is legible' you mean only that you can read it: you do not know whether I can (McLean 1980)"

Legibility has frequently been explained as a complex interaction of multiple factors dependant on the design of the message (typeface, print design, layout), surface (matte, glossy), viewing conditions (light levels) and reader characteristics (visual acuity, color perception) (Pyke 1926; Tinker 1963; Bix 2003). Legibility has also been named very differently in different sources. Some terms commonly used include legibility, reading ease, perceptibility, encoding ease etc. This document refers to all the above as simply, 'legibility' for the purpose of clarity.

However, several authors agree that legibility is best defined as the ability of readers to read the printed message with ease (Tinker 1963; Pietrowski 1993; Bix 1998; Bix 2001). There are many factors that affect the ability of readers to read printed messages. Three of the major factors are described in the following section.

Factors that affect legibility:

As mentioned previously, optimum legibility is the result of several factors and seemingly disparate systems working together (Tinker 1963; Pietrowski 1993; Bix 1998; Bix 2001).

Typeface and Message Design: Typographic design has an important bearing on the visual impact and reading ease of printed matter. The morphology of a typeface includes terms like: x-height, serifs, ascenders, descenders, uppercase, lowercase letters, kerning etc (McLean 1980; Bix 2001). Readers are directed to the aforementioned references for visual representations of these factors on popular typefaces. Each one of these factors can impact the legibility of text. Reading ease is also affected by the size and width of the message (Tinker 1963; McLean 1980; Bix 2001).

Other factors that impact the legibility of printed matter are characteristics of the message itself, like line length, width and line-spacing (leading).

Printing Surface: In his widely cited report to the H.M.Stationary office in 1926, Pyke summarizes several characteristics that printing paper should have in order to have maximum legibility (Pyke 1926). The general recommendations were that printing paper should have very little specular reflectance (gloss) and a high degree of diffuse reflectance, in other words, a matte surface. From a paper-making perspective, the surface of the paper depends on several factors like pulping process, additives and coatings. These coatings and additives may increase the surface smoothness and create higher levels of specular reflectance. There are also specific recommendations regarding the opacity and thickness of the paper for optimum print legibility.

Contrast: Contrast here, refers specifically to the difference in color between the print and the background. This has been found to impact print legibility to a great extent. Several studies have shown that dark text on light or neutral backgrounds have high levels of legibility (Tinker 1963; Bix 1998; Bix 2001).

FDA's labeling regulations for OTC drug packages recommend that *"the type must be all black or one dark color, printed on white or other light, neutral color, contrasting background (FDA 1999)"*. This regulation is consistent with the concepts mentioned above.

Additionally, Tinker mentions in his book that the *"use of white print on black background is attention getting to a striking degree...(Tinker 1963)"*

Measurement Techniques:

Since legibility is a very complex concept resulting from the interaction of several print and reader characteristics, measuring it is challenging, to say the least. Nevertheless, there have been several measurement techniques that have been proposed as measures of print legibility over the years. Some of these are explained below. A bulk of the information presented here draws from a book chapter in the 1963 book, Legibility of Print by Tinker (Tinker 1963).

Visibility Methods: These methods use photographic filters positioned like eye glasses in front of the eyes. These filters are rotated in order to increase and decrease the apparent brightness of the print sample that is being viewed. Based on the extent of rotation of these filters, legibility has been measured. This principle has been used in the design of the Luckiesh-Moss Visibility meter (Luckiesh and Moss 1938; Luckiesh and Moss 1939).

Distance methods: In distance methods, print samples are placed at different distances from the reader's eyes and the relative ease of reading between the different print samples are evaluated (Sanford 1888; Tinker 1963).

Quick-Exposure Methods: Legibility is also measured by evaluating the speed at which readers can read printed information. Previous experiments have been performed by exposing test subjects to printed matter for very short amounts of time through use of instruments like the Tachistoscope (Dockeray 1910). A tachistocope is an optical device that allows researchers to quickly present stimuli to subjects. It flashes images at a predetermined rate. These were essentially precursors to modern setups that use computer software and projectors. Similar techniques include speed-reading measurements where reading speed has been linked to legibility.

Focal Variator Methods: These are similar to the visibility meters, but they use lenses instead of photographic filters. The rotation of the lenses can increase or decrease the blur of the printed matter that is being viewed. Legibility is calculated based on how soon a reader can recognize and read printed matter (Weiss 1917).

Eye Based Methods: Eye movements and blinking rate have also been used to measure legibility. Researchers have performed direct counting and recording of blinks and have tried to correlate this with legibility. The available literature indicates that illumination, subject position and other ambient factors must be kept constant during this whole process (Tinker 1946).

Chapter 4 on Eye tracking mentions cases where eye tracking has been used to study reading behavior. Such measures have also been used in the past to evaluate efficient message design and print design. These could be used to evaluate print legibility as well (Rayner 1992).

Legibility Measurement in this study:

This study makes use of the Lockhart Legibility Instrument (LLI) which was developed at the School of Packaging, Michigan State University (Bix 2001). Detailed diagrams and the working of this instrument shall be presented in the Materials and Instrumentation chapter.

The ASTM standard for comparative legibility now uses the LLI and this method will be used in this study to evaluate the legibility of the various color combinations of the PWL's (ASTM 2006).

Chapter 6: Materials and Instrumentation

The first part of this chapter features an in-depth discussion of the materials and supplies used in this study. The second part includes detailed descriptions of the instruments used and their working.

Materials

Near Point Visual Acuity card: This is a laminated card which is printed with text of gradually increasing type sizes from bottom to top (See Figure 5). The subject is asked to hold this card at around 16” from their eyes. They are then asked to read the lowest line possible. The subject’s visual acuity is determined based on the lowest line that is read with ease. This visual acuity card is manufactured by Dow Corning® Ophthalmics and the visual acuities that can be measured range from 20/20 to 20/120.

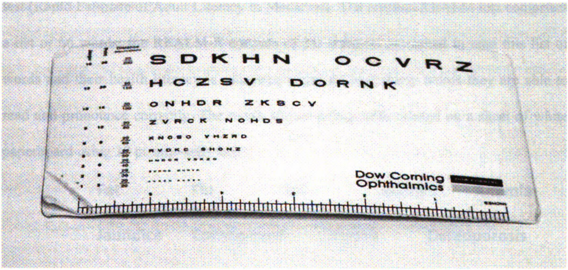


Figure 5 - The Dow Corning Opthalmics, Near Point Visual Acuity Card

REALM-R Card: *"Low literacy is a pervasive and under recognized problem in health care. Approximately 21% of American adults are functionally illiterate, and another 27% have marginal literacy skills. Such patients may have difficulty reading and understanding discharge instructions, medication labels, patient education materials, consent forms, or health surveys." (Davis, Michielutte et al. 1998)*

The above quote from an article by Davis et al. illustrates the important of health literacy in today's world of self-medication and out-patient care. It was decided to include a test of health literacy along with the eye tracking study in order to account for the effect of low health literacy on the noticeability and recall of warning label information. For instance, subjects who did not recognize the terms used in the warning messages might spend more time on them while examining the vials.

The Realm-R (Bass, Wilson et al. 2003) is a shortened version of the REALM test (Rapid Estimate of Adult Literacy in Medicine). The original REALM test comprises a list of 66 words; the REALM-R consists of 11. Subjects are asked to read this list of words and their health literacy is estimated based on how many words they are able to read and pronounce correctly. The words shown below were printed on a sheet of white paperboard using 20 pt sans-serif font.

Fat **Flu** **Pill** **Allergic** **Anemia**

Jaundice **Constipated** **Colitis** **Osteoporosis**

Fatigue **Directed**

Subjects were asked to read the words and say "blank" when they came to a word they could not read. Subjects are scored based on their ability to correctly pronounce the words. However, the words fat, flu and pill are not scored, serving as an acclimation period. Subjects with a score of 6 or less are considered to be at risk for poor health literacy.

Pseudo Iso-Chromatic Plates: A booklet with pseudo Iso-chromatic plates manufactured by Richmond products (Albuquerque, New Mexico) was used to quickly screen subjects for red/green color blindness. The subjects were asked to flip the pages and read the numbers (see Figure 6). Out of the 14 images, subjects are considered to have normal color vision if they were able to read 10 or more.

Figure 6 - Pseudo Iso Chromatic Plates booklet used for testing for Red/Green Color blindness

Vials: The vials used in this study were 10 dram vials of the 1-Clic type (Rexam Prescription, Ohio). Previous studies have shown that vials of the 10 dram size are the most commonly used for dispensing prescription medication in the United States (De La Fuente 2006). Vials were obtained through the campus pharmacy (see Figure 7).

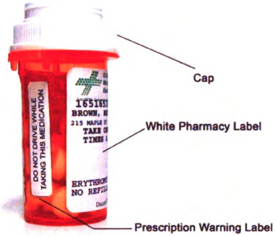


Figure 7 - A photograph of a vial used in the study. A common 10 dram prescription vial of the 1 Clic type with two types of labels; PWL's and White Pharmacy Labels

Labels: The Five color contrast combinations of PWL's that were tested in this study were,

- Black text on Blue background
- Black text on Red background
- Black text on Yellow background
- Black text on White background
- Blue text on White background

The first four color combinations in the above list were chosen from pharmaceutical supply catalogues as the most commonly commercially used combinations (Pharmex 2009). A fifth combination (Blue text on White background) was added based on the expert recommendation of Mr. Tom Ball at the Olin Health Center, MSU. At the University pharmacy, the light blue on white contrast is being used on PWL's for vials used to dispense tablets for typhoid. They carry instructions that direct patients to refrigerate the vials. Pharmacists indicated that many patients had failed to do so, and suggested that this contrast be included in the study.

Labels were designed using Adobe® Illustrator® CS3 (See Figure 8) and were printed using high fidelity digital printing. The labels measured, 4 centimeters by 1 centimeter and were pasted to the vials as shown in the Figure 7. For the eye tracking portion of the study, the vials were placed inside pharmacy bags and handed to subjects.

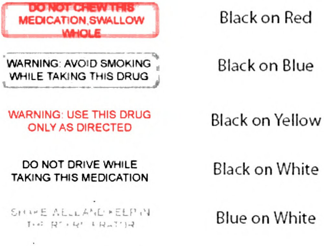


Figure 8 - PWL's used in this study

In addition to the PWL. each vial also had a White Pharmacy Label affixed as shown in Figure 7. This label contained information about the pharmacy, dosage, patient, doctor and drug. Figure 9 depicts one of the prescription labels used in this study.

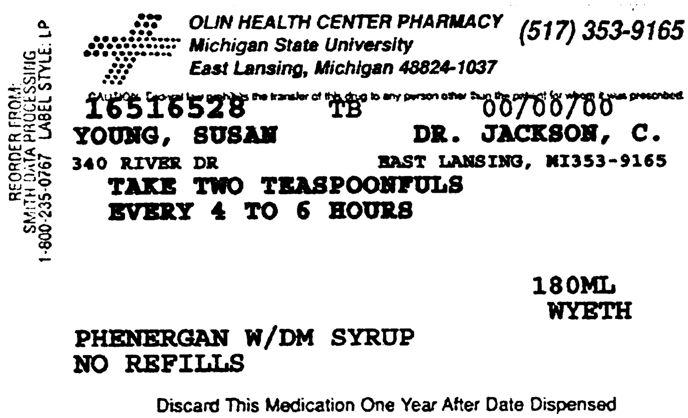


Figure 9 – An example of a White Pharmacy Label from the Olin Health Center that was pasted on all vials.

Warning Label Text: It has been mentioned that the composition of the text that a person reads strongly affects the gaze trail. It was recognized early in the study that the five pieces of warning label text used as stimulus material had to be scientifically evaluated and standardized for reading ease.

The Flesch-Kincaid reading-ease score, introduced by Rudolph Flesch in 1948 is widely used in general readability assessment (Flesch 1948) and comes bundled with popular word processing packages such as Microsoft® Word. The reading ease score of a piece of text can be calculated in Microsoft® Word by it and clicking on Word Count in the Tools Menu. This method of readability evaluation was used to create five warning messages that closely matched those used commercially for PWL's. The warning messages used on PWL's are shown in Table 2.

A Flesch-Kincaid reading ease score of 66.7 falls in the range that is easily readable by most 13-15 year old students (Dubay 2004).

Table 2 - Warning message readability scores

PWL	Warning label Text	Flesch-Kincaid Readability Score
Black on Red	Do not chew this medication, swallow whole	66.7
Black on Blue	Warning: Avoid smoking while taking this drug	66.7
Black on Yellow	Warning: Use this drug only as directed	66.7
Black on White	Do not drive while taking this medication	66.7
Blue on White	Shake well and keep in the refrigerator	66.7

Instrumentation:

The ASL 501 Head Mounted Optics Eye Tracker

An Applied Science Laboratories (ASL) (Boston, MA), bright pupil eye tracker (ASL 501) was used in this study (see Figure 10). The construction and working principle of this instrument are described below.

Construction: The head mounted optics eye tracker consists of a headset that houses two cameras and a monocle arrangement. The 'Eye Camera' is used to record images of the eye and the 'Scene Camera', as the name suggests, is used to record the subject's field of view. The eye camera is coaxial to a lamp that shines a near infrared beam of light onto the monocle. The monocle is coated with a special coating that reflects the light beam. This beam falls on the eye and illuminates the pupil, thus the name 'Bright Pupil'. A part of this light beam is reflected off the surface of the cornea and forms a small bright spot of light known as the Corneal Reflection or CR. The positions of this CR spot and the large bright pupil depend on the direction of gaze of the eye.

Figure 10 - ASL 501 Head Mounted Optics Eye Tracker

Calibration: The floating calibration procedure for the eye tracking system involves directing subjects to look at 9 different points on a single plane in the field of view. The positions of the CR and pupil are recorded at each of the nine points during calibration. Based on the gaze information for these nine calibration points, the recording software extrapolates the point of fixation of the eye for all points in the visual field.

Setup: As such, the ASL 501 requires that the subject's head be stationary at all times. In order to calibrate as accurately and precisely as possible, a custom calibration system was built, which incorporates both a pane of glass and a chin rest (see Figure 10).

Figure 11 - Custom Designed Eye Tracking Table with Glass pane and Chin rest arrangement

Data Recording: As mentioned, the eye camera records the position of the eye. The scene camera records a video of the subject's field of vision. The recording software, Eye Response® Gaze Tracker® (Boston, MA), calculates the points of fixation and superimposes a set of crosshairs over the scene video.

A short note on the process of analyzing eye tracking data will be provided in the Methods chapter.

Legibility Instrument:

The Lockhart Legibility Instrument (LLI) was developed by MSU researcher, Hugh Lockhart, and is used to quantify the relative legibility of printed matter. It is mentioned in the literature that the LLI was initially used as a tool to study noticeability

(Bix 2001). However, the instrument was later redesigned and its methods reviewed to enable objective measurement of legibility (Bix 2001). In this study, the LLI was used to assess the comparative legibility of the five color contrasts in accordance with the methodology specified by ASTM D7298 (ASTM 2006).

Construction: The LLI is a large viewing box that is 50" by 19" by 19". A movable easel inside the instrument is illuminated by two 25W incandescent flood lights. The amount of current fed into these lamps is regulated by a rheostat actuated by a knob. The light level is measured by a light meter built into the easel. The light meter displays its output on a digital display mounted on the body of the LLI (see Figure 12).

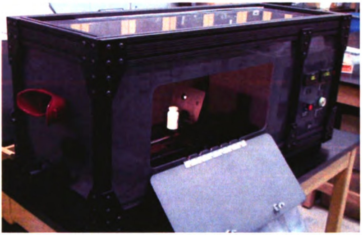


Figure 12 - Lockhart Legibility Instrument

Operation: Subjects view the easel through a viewer that directs their gaze into the instrument through a set of polarizing filters. At their initial position, the filters polarize all light that passes through them and the inside of the box appears completely dark to the subject. As the subject rotates a flywheel mounted below the viewer, one of the polarizing filters rotates and allows light to pass through. It, thus, appears to the subject that the label on the easel is progressively receiving more light. The subject is instructed to turn the wheel until the first point where they can read the text on the label with ease.

The angle through which the polarizing filter has rotated is known as a legibility index (0-90°) and it represents the relative ease of reading of the printed message. The legibility index values range from 0 to 90, with 90 being most difficult to read. This legibility index reading is taken down from a digital display on the LLI. The LLI used in this study was manufactured by Sycamore Industrial Services, Holt, MI.

Chapter 7: Methods

Working Hypotheses:

There were two working hypotheses while studying the PWL's in this study.

- The black on red color contrast combination is more noticeable than the black on white, black on blue, black on yellow and blue on white color contrast combinations
- The black on white color contrast combination is more legible than the black on red, black on blue, black on yellow and blue on white color contrast combinations.

Test Procedure:

Individual subjects were tested using the following procedure.

Consent:

Researchers provided subjects with a verbal explanation and an overview of all test procedures. All test subjects were shown a consent form which they were asked to review and sign. A copy of this form is provided in the Appendix 1.

Eye Tracking:

A random order of presentation of vials was determined for each subject using a custom written ActionScript® randomizer program on Adobe Flash CS 3®.

Subjects were seated in a chair with adjustable height at the eye tracking table that was fixtured with a pane of glass and a chin rest. The head band of the ASL 501 eye tracker was placed on the subject's head and tightened in place while making sure that they were in a comfortable position. Minor adjustments to the optics system and the chin rest were made so that they were appropriate for the subject. Once this was done, subjects were asked to rest their chin on the chin rest.

Researchers then asked the subject to hold as still as possible, while they performed the calibration procedure using the floating calibration technique described in the previous chapter. This technique calibrated the subjects to the pane of glass that is permanently affixed to the desk. Fixing the distance between the subject and the viewing field allows the eye tracker to precisely track the position of the subject's eye in space and minimizes the potential for parallax error.

Each treatment (the five vials with warnings in different color contrast combinations (see Figure 7) was placed in a bag from the pharmacy, and randomly handed to the subjects who were instructed to remove the vial from the bag and examine it as they pleased, for as long as they wished. Subjects were provided with the following scenario,

"You have been delivered new prescription drugs from the pharmacy. Please do what you would normally do." Researchers also instructed the subjects as follows. "If you read anything on the vial, please press the vial against the glass while doing so."

The eye tracker system software outputs a video file with the gaze trail of the subject. This allowed researchers to determine whether or not the subject noticed the PWL. It also allowed researchers to quantify the time that subjects spent on each of the label and package elements.

A control population of 15 subjects (aged 18-29) and a test population of 17 subjects (aged 50 +) viewed each of the five labels using the Applied Science Laboratories (ASL) eye tracking device. As a precaution, a greater number of subjects were tested in the older population to account for the possibility of unusable eye tracking recordings.

Recall and Recognition Test:

After testing with the eye tracking equipment, subjects were given a plain sheet of paper and a pen and were asked to recall everything they could about the prescription vials that they just viewed. This was a test of recall. A recognition test was then administered in which subjects were shown a sheet of paper with 10 labels, in 10 different color contrasts (See Figure 13).

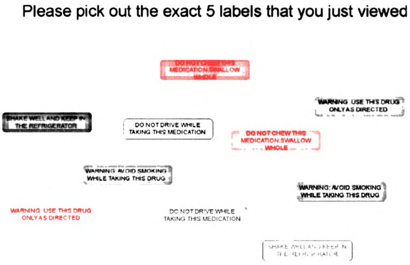


Figure 12 - Recognition Sheet

Five of these were identical to the labels on the vials that subjects had just viewed using the eye tracker. The other five had the same messages, in different color contrasts (black on purple, black on orange, black on green, black on pink and dark green on light green). Subjects were asked to pick out the exact five labels that they had just viewed.

Collection of subject information:

Data recording sheets were used to gather subject information such as age, ethnicity, gender, educational background and number of prescription drugs taken per day. Subjects were assigned a numeric code in order to provide strict anonymity. They were referred to by this number throughout the data processing and analysis stage. A sample data recording sheet is provided in Appendix 3.

Additional Testing:

Health Literacy Testing: The REALM-R card was held at a convenient reading distance from the subject. Subjects were provided with the following verbal instructions.

"Sometimes in the health care system, medical words are used that many people are not familiar with. I would like to get an idea of what medical words you are familiar with. I would like you to read the words printed on this card aloud. When you have trouble with a word, please say 'Blank' and move on to the next word"

Checklists were used to record the number of words correctly pronounced. If a subject took more than five seconds to read a word, that word was considered a failure. A score of lower than six indicated that subjects were at risk for poor health literacy. At risk denotes the fact that patients may not be able to understand the words that they are expected to while interacting with a physician in a primary care setting.

Color Perception: To examine color perception and screen subjects for red/green color blindness, subjects were asked to read the numbers from a booklet with a set of 15 pseudo-isochromatic plates manufactured by Richmond products. Subject responses were recorded and tabulated.

Visual Acuity: The visual acuity of the subject was then evaluated using the Dow Corning ® Opthalmics Near Point visual acuity card. Subjects were asked to wear their glasses or lenses, bifocal or contact while testing for near vision. The card was held at 16 inches from their eyes. The visual acuities of subjects were recorded and tabulated based on the lowest line that subjects could read from the card

Legibility testing:

Subjects viewed the fifteen PWL's in three sets (see Table 3 for example) using the Lockhart Legibility instrument, each set containing the five PWL color contrast combinations in random order. At the end of testing, the legibility index values were averaged across the three sets. This provided researchers with 33 replicates for each label.

For each contrast combination, the subjects were instructed to turn the wheel on the LLI until they were able to read the text on the labels with ease. The angle through which the wheel turns is known as the legibility index and this provides a measure of the relative legibility of the piece of printed matter. All testing was done in accordance with ASTM standard D7298-06, Standard Test Method for Measurement of Comparative Legibility by Means of a Polarizing Filter Instrumentation (ASTM 2006).

It is known from previous studies that while testing on the LLI, the initial readings differ from later ones (Bix 1998; Bix 2001; Bix 2003). This necessitated the introduction of two dummy labels prior to the actual testing of the PWL's. Their legibility scores were not considered for analysis. Previous studies performed using this instrument used an easel distance of 17.5" and a light level of 25 foot candles. This study also used the same settings (Bix 1998; Bix 2001; Bix 2003).

	Dummy Set	Set 1	Set 2	Set 3
Subject 1	Two Dummy Labels were tested at the beginning of the Test	Black on Red	Black on White	Black on Yellow
		Black on Yellow	Black on Red	Black on Blue
		Black on White	Blue on White	Black on White
		Black on Blue	Black on Yellow	Blue on White
		Black on Blue	Black on Blue	Black on Red

Table 3 - Example legibility test order

Compensation:

Once testing on the LLI was completed, each subject was given a cash reward of 30$. Subjects were thoroughly debriefed.

Data Processing:

Eye tracking: The gaze trail videos obtained using the eye tracker were analyzed using the Eye Response® Gaze Tracker® analysis software. The images of the vials in the videos were broken down into areas known as "look zones". The amount of time the subjects spent on each of these look zones was calculated using the software. The specific areas of the vial that were studied were:

- The White Pharmacy Label
- The Cap
- The PWL

Other important information, such as the number of times that a look zone was entered (hit count) and the total time subjects spent examining each vial was also recorded.

Legibility: Legibility indices or degrees of rotation of the polarizing filter were coded in a spreadsheet that included subject demographics and characteristics.

Results, Analysis & Discussion

Data Encoding and Entry:

As explained in the Materials and Methods chapters, there were several different forms of data that were collected during the course of this study. Eye tracking data was collected as video files that were later processed using the Gaze Tracker® analysis software. The result of this analysis was a Microsoft Excel® spreadsheet which had zone-specific metrics. Legibility data (degrees of rotation of the polarizing filter) and subject characteristics (visual acuity, health literacy etc) were coded and manually entered into spreadsheets. These items were bundled into a large spreadsheet in the 'flat file' form suitable for entry into the SAS® Statistical Software.

This chapter includes summaries of data sets from this study and subsequent mixed model statistical analyses using the SAS® Statistical Software.

Subject Demographics:

The control population was comprised of 15 subjects between the ages of 18 and 29 years (ave.22.8 ± 2.33). The test population was made up of 17 subjects who were over 50 years of age (ave. 62.2 ± 8.10. Demographics of these two groups are provided in the table below. Subjects ranged in age from 20 - 77 years.

The following tabulations and corresponding charts depict the distribution of subject gender and reported ethnicity (see Table 4), health literacy and Red/Green color blindness (see Table 5), prescription drugs per week (Table 6). Graphs are shown after the respective tables to represent this information in a visual manner.

Table 4 – Subject Gender and Ethnicity (Data for Figure 14)

Popln	Gender					Ethnicity			
	# of Male Subjects	Avg Age ± SD	# of Female Subjects	Average Age ±SD	Total # of Subjects	# of Caucasian Subjects	Avg Age ± SD	# of Asian American Subjects	Avg Age ±SD
Control Population (18-29)	8	22 ± 1.06	7	23.7 ± 3.09	15	10	22.9 ± 2.42	5	22.6 ± 2.40
Subject Population (50+)	5	65.4 ± 6.54	12	61 ± 8.59	17	17	62.2 ± 8.10	0	0
Total	13	38.6 ± 22.31	19	47.2 ± 19.74	32	27	47.7 ± 20.45	5	22.6 ± 2.40

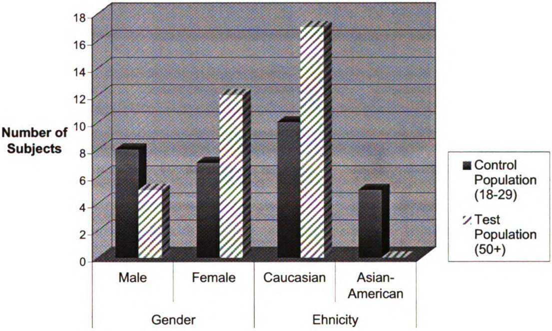


Figure 14 – Subject Gender and Ethnicity

Table 5 – Health Literacy and Red/Green Color Blindness (Data for Figure 15)

Population	Health Literacy		Color Blindness	
	# of individuals Not at risk	# of individuals At Risk	# of individuals Not at risk	# of individuals At risk
Control Population (18-29)	11	4	15	0
Subject Population (50+)	16	1	15	2
Total	27	5	30	2

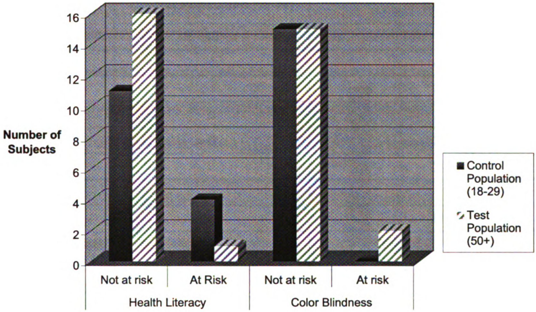


Figure 15 - Health Literacy and Red/Green color Blindness

Only one member of the test population tested to be at risk for low health literacy, in contrast to 4 members from the control population (aged between 18 and 29 years). The native language for all subjects from both populations was English. Only one subject out of the 32 that were tested had less than a college degree. All other subjects ranged in their level of education from at least two years of college to a doctoral degree. All members of the control population tested as having normal red/green color vision while 2 members of the test population tested to be at risk for abnormal red/green color vision.

The data on Table 6 and Figure 16 indicates that a majority of the subjects in the control population (18-29 years of age) did not report taking any prescription drugs on a daily basis at the time of the study. However, a majority of the test population (50+) reported taking more than 2 prescription medications every day.

Table 6 – Medication Consumption (Data for Figure 16)

Number of Prescription Drugs Per Day	# of Subjects within the Control Population (18-29)	# of Subjects within the test population (50+)
None	10	1
1	4	4
2	1	3
3	0	3
4	0	3
5	0	2
Over 5	0	1
Total	15	17
Average # of daily prescriptions ± SD	0.4 ± 0.63	2.94 ± 2.16

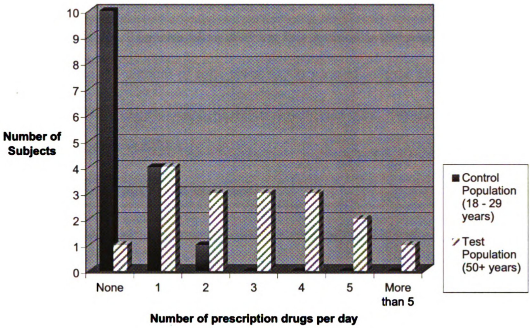


Figure 16 - Medication Consumption

Noticeability of PWL's using Eye Tracking data:

Eye tracking data was collected in the form of video files and these files were analyzed by zone (PWL, White pharmacy label, Cap - see Figure 7). Dependent variables for analysis of the eye tracking data were both binary and variable:

- The proportion of subjects who noticed a zone (yes/no; a binary attribute)
- The proportion of total time per vial on a zone, i.e. Time spent on a particular zone divided by the total time spent examining the vial itself (continuous variable)
- Number of times the zone was crossed by the subject's gaze trail for each zone of each vial (number of hits) (discrete variable)

Previous studies have quantified noticeability based on several different dependant variables. The time taken until the first fixation in the zone of interest, number of fixations and the amount of time spent in the zone of interest have been used in the past (Krugman, Fox et al. 1994). In this study noticeability was defined and quantified using three response variables.

- The probability of noticing a zone
- The proportion of total time spent on a zone standardized by zone density and
- The total number of hits on each zone of each vial by a subject

Further explanations and statistical analyses for each of the three dependant variables are provided below.

The probability of noticing a zone:

The success/failure of a subject to notice a zone (cross into a zone in their gaze trail) was modeled as a binary response variable (noticed / did not notice) with zone being a seven level categorical variable. The seven levels were: the white pharmacy label, the five PWL's and the cap. This data was available for 32 subjects (15 controls, 17 test subjects). The frequency distribution table of the number of subjects who noticed/did not notice label zones is shown below (see Table 7). As can be seen from the table, all subjects (100%) noticed the White Pharmacy label while at least 10 subjects failed to notice each of the PWL labels.

Table 7 - Frequency Distribution Table of Noticeability Binary Variable

Noticed (yes/no)		PWL					Cap	White Pharmacy Label
		Black on Red	Black on Blue	Blue on White	Black on Yellow	Black on White		
No	Frequency	11	11	10	11	11	124	0
	Percent (%)	34.375	34.375	31.25	34.375	34.375	77.5	0
Yes	Frequency	21	21	22	21	21	36	160
	Percent (%)	65.625	65.625	68.75	65.625	65.625	22.5	100
Total number of subjects	Frequency	32	32	32	32	32	160	160
	Percent (%)	100	100	100	100	100	100	100

This probability of noticing a zone was modeled as a function of population (control versus test), zone (PWL's - in 5 contrasts, cap and pharmacy label), their 2-way interaction, gender and total time spent on the vial. The model was fitted using a logistic regression model with the GLIMMIX procedure of SAS (Version 9.1, SAS Institute Inc., Cary, NC). The effects of health literacy, number of prescription drugs per day, ethnicity, age, color blindness, order of presentation of vials and visual acuity of the subject were included in the statistical model during the initial stages of analysis, but were later dropped based on the type III test p-values ($p > 0.10$). Pairwise comparisons were conducted using Bonferroni adjustment to avoid inflation of Type I error rate.

An effect of label zone was identified on the probability of noticing it ($p < 0.0001$). Pair wise comparisons indicated that there was a significantly greater probability of noticing the White Pharmacy Label compared to any PWL, regardless of color ($p < 0.01$ for all corresponding pair wise comparisons) and a significantly lower probability of noticing the cap compared to any PWL regardless of color ($p < 0.01$ for all corresponding pair wise comparisons - see Figure 17). There was not enough evidence to conclude an effect of color contrast on the probability of noticing a specific PWL ($p > 0.60$ for all corresponding pair wise comparisons). The data did not support a difference in the dependant variable between control and test population results ($p = 0.98$). Averaged across label zones it was also identified that females had a greater probability of noticing any zone as compared to the males ($p < 0.0001$).

Table 8 - Predicted Probability of Noticing a Zone (Data for Figure 17)

Zone		Probability of Noticing a Zone (%)	Standard Error
White Pharmacy Label		100	0.0011
PWL	Blue on White	69.62	8.998
	Black on Yellow	67.14	9.375
	Black on Red	66.91	9.382
	Black on White	66.63	9.421
	Black on Blue	65.2	9.132
Cap		17.11	3.301

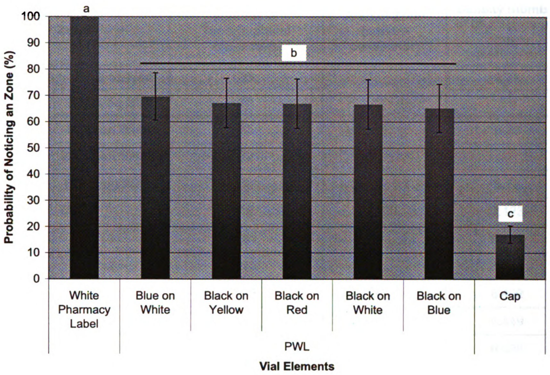


Figure 17 - Predicted probability of noticing a zone ± estimated standard error. Letters (a, b, c) indicate statistically significant differences ($p < 0.01$)

The Proportion of total time spent per zone standardized by zone density:

The proportion of the total time spent per zone per vial (time in a zone/total time per vial) was calculated during analysis with the Gaze Tracker® Analysis software. Recognizing the fact that different zones had differing amounts of text, and therefore, would require varying times, researchers attempted to standardize the proportion of time by dividing by the "zone density." The zone density was calculated by dividing the number of typographical characters in each zone by the area (see Table 9 below).

Table 9 - Zone density values for all label zones

Zone		Number of Typographical characters	Area of Zone (square centimeter)	Density (number of characters per square centimeter)
PWL's	Black on Red	35	4	8.75
	Black on Blue	38		9.5
	Blue on White	33		8.25
	Black on Yellow	32		8
	Black on White	35		8.75
Cap		15	7.06	2.123
White Pharmacy Labels	1	261	42.64	6.121
	2	270		6.332
	3	273		6.402
	4	275		6.449
	5	277		6.496

This dependant variable (standardized time, or proportion of time/zone density) was analyzed and it was revealed that a large number of subjects had gaze trails that did not cross into several PWL's and caps. This created several zeroes in the pool of dependant variable values, causing a problem in the residual plots that were generated.

Subsequently, this data set was truncated to include only those values that were not zero, i.e., those instances where subjects had not logged into a given label zone were eliminated from analysis.

After this, the response variable "standardized proportion of time" was log transformed and modeled using a general linear mixed model fitted with the MIXED procedure of SAS (Version 9.1, SAS Institute Inc., Cary, NC). The model included the fixed effects of treatment, population (control vs test), total time spent on the vial and all

2-way interactions. Also, the model included the random effect of subject nested within population in order to account for the split-plot structure of the data.

The effects of health literacy, number of prescription drugs per day, ethnicity, age, gender, order of presentation and visual acuity were considered for model inclusion but did not make a significant contribution to model fit based on Bayesian Information Criteria (BIC). Thus, these explanatory variables were not included in the final model. However, the fixed effect of color blindness was included in the final model based on BIC.

Pairwise comparisons were conducted on the estimated least square means of the dependant variable using Bonferroni's adjustment to avoid inflation of Type I error rate. Model assumptions were checked and considered to be appropriately met. The salient findings from this analysis are provided below.

The evidence did not support an effect of population (control vs test) on the standardized proportion of time spent on each zone. ($P = 0.3453$). After adjusting for the effects of population and treatment, an association was identified between color blindness and standardized proportion of time ($P = 0.0209$). Subjects that failed the red/green color blindness test had larger standardized proportion of time across all zones regardless of population (control vs test) compared to those that passed the color blindness test.

During the analysis, a significant interaction effect between zone and total time spent on the vial was identified on the standardized proportion of time. As a result, the main effect of zone on the standardized proportion of time should not be interpreted on its own. Thus the dependant variable estimates were calculated at four values of total time (10 seconds, 20 seconds, 30 seconds and 40 seconds)

The tabulation (see Table 10) provides the Bonferroni adjusted pair wise comparisons for the estimated least square means of the dependant variable at each of the four values of total time. In the table, the "standardized proportion of time" dependant variable is not statistically different for zones that share the same letter designation. It can be seen that in all groups, there was not enough evidence to conclude an effect of color contrast of PWL on the standardized proportion of time spent (all PWL zones share the same alphabet). At the same time, the proportion of total time standardized by zone density for the white pharmacy label was higher than that of the PWL and the cap. In other words, subjects spent more of the total time on the white pharmacy label regardless of differences in density and area.

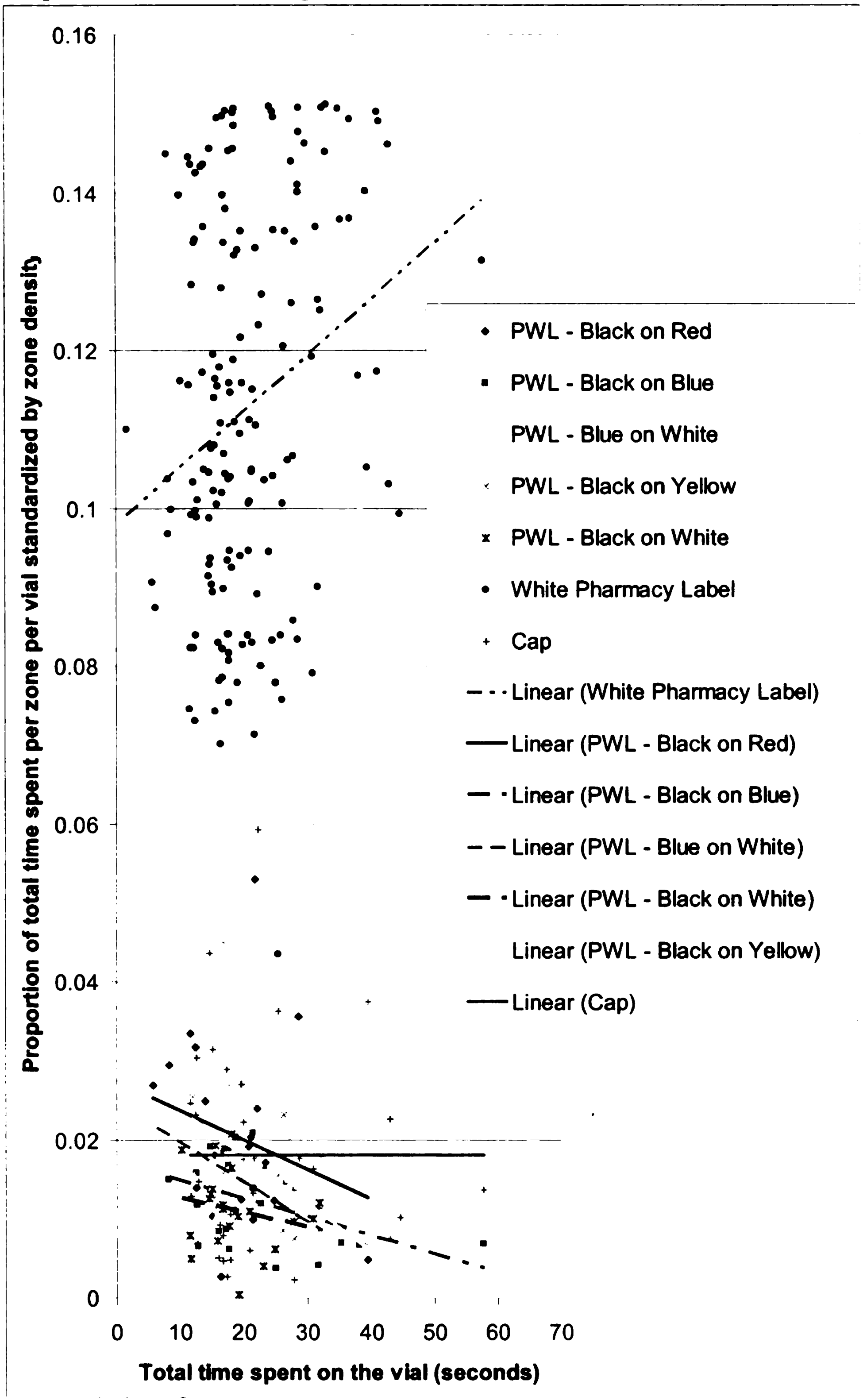
Table 10 - Bonferroni adjusted pair wise comparisons across zones for four levels of total time spent per zone.

Total time spent per vial			
Total time spent on the vial is 10 seconds		Total time spent on the vial is 20 seconds	
Zone	Pair wise Comparison	Zone	Pair wise Comparison
Black on Red	A	Black on Red	D
Black on Blue	A	Black on Blue	D
Blue on White	A	Blue on White	D
Black on Yellow	A	Black on Yellow	D
Black on White	A	Black on White	D
White Pharmacy Label	B	White Pharmacy Label	E
Cap	A	Cap	D
Total time spent on the vial is 30 seconds		Total time spent on the vial is 40 seconds	
Zone	Pair wise comparison	Zone	Pair wise comparison
Black on Red	G	Black on Red	J
Black on Blue	G	Black on Blue	J
Blue on White	G	Blue on White	J
Black on Yellow	G	Black on Yellow	J
Black on White	G	Black on White	J
White Pharmacy Label	H	White Pharmacy Label	K
Cap	G	Cap	J

The graph on the following page plots the value of total time spent examining the vial (x-axis) against the proportion of total time standardized by zone density for each zone (y-axis) for both subjects of both populations. Seven scatter plots were created on this graph, one for each zone (5 color contrasts of PWL's, the white pharmacy label and the cap). Each scatter plot was used to generate a colored trend line. The figure has seven trend lines that provide an overview of how subjects spent their time on zones as the amount of total time they spent examining the vials increased.

As observed in the graph and as revealed in the pair wise comparisons shown above, subjects, on the whole, spent a lot more time in the white pharmacy label zone as compared to the other zones. The graph also seems to indicate that as the amount of time spent on the whole vial increased, subjects spent more time on the white pharmacy label. It is speculated that when subjects spent a large amount of time examining the vial, they were spending a greater proportion of their time reading the white pharmacy label and trying to glean as much information as they could from it. Despite spending larger amounts of time on the vial, subjects spent a smaller proportion of their total time on the PWL.

Figure 18 - Scatter Plots and Trend Lines relating Total Time spent on vials to the Proportion of total time spent on each zone standardized by zone density



The total number of hits on a zone:

The total number of hits was recorded for the 3 zones on each vial: namely the white prescription label, the PWLs (5 contrasts) and the cap. Similar to the previous variables, zone was defined to be a categorical variable of 7 levels: the White pharmacy label, 5 color contrasts of PWL's, and the cap. This data is shown below in the form of a cumulative frequency table with the number of hits for each zone ranging from 0 to 4 (see Table 11).

Table 11 - Cumulative frequency table for the number of hits on each label zone

Number of Hits		PWL					Cap	White Ph'cy Label
		Black on Red	Black on Blue	Blue on White	Black on Yellow	Black on White		
0	Freq	11	11	10	11	11	124	0
	Percent (%)	34.375	34.375	31.25	34.375	34.375	77.5	0
1	Freq	19	18	20	18	20	33	96
	Percent (%)	59.375	56.25	62.5	56.25	62.5	20.625	60
2	Freq	2	3	2	3	1	3	58
	Percent (%)	6.25	9.375	6.25	9.375	3.125	1.875	36.25
3	Freq	0	0	0	0	0	0	4
	Percent (%)	0	0	0	0	0	0	2.5
4	Freq	0	0	0	0	0	0	2
	Percent (%)	0	0	0	0	0	0	1.25
Total Number of subjects	Freq	32	32	32	32	32	160	160
	Percent (%)	100	100	100	100	100	100	100

This 'total number of hits' response variable was modeled as Poisson distributed using a generalized linear mixed model. This model included the fixed effects of zone, population (control and test), their 2-way interaction and the effects of gender and total time spent on the vial. Also, the model included the random effect of subject nested within population in order to account for the split-plot structure of the data. The model was fitted using the GLIMMIX procedure of SAS (Version 9.1, SAS Institute Inc., Cary, NC). The effects of health literacy, number of prescription drugs per day, ethnicity, age, color blindness, order and visual acuity were considered but were not included in the final model based on type III test p-values ($p > 0.10$).

Table 12 - Expected number of hits on a zone (Data for figure 19)

Zone	Predicted Mean Number of Hits	Standard Error
White Pharmacy Label	1.3585	0.1101
Black on Red	0.7122	0.1496
Black on Yellow	0.6996	0.1472
Black on White	0.6712	0.1433
Blue on White	0.6645	0.1453
Black on Blue	0.6519	0.1447
Cap	0.2003	0.03793

The number of hits, as obtained from the gaze trail analysis differed by zone ($p < 0.0001$). There were a significantly greater number of hits on the White pharmacy label when compared to the PWL, regardless of PWL color contrast ($p < 0.05$ for all corresponding Tukey adjusted pair wise comparisons - See Table 12 and Figure 19). The number of hits on the cap was less than all PWL's regardless of their color ($p < 0.01$). The evidence did not support an effect of color contrast on the number of hits on PWL's ($p > 0.99$ for all pair wise comparisons) (see Table 12 and Figure 19).

The estimated mean number of hits and corresponding standard error are provided for each label zone (see Table 12 and Figure 19). Pair wise comparisons were conducted using the Tukey Kramer adjustment to avoid inflation of the Type 1 error rate.

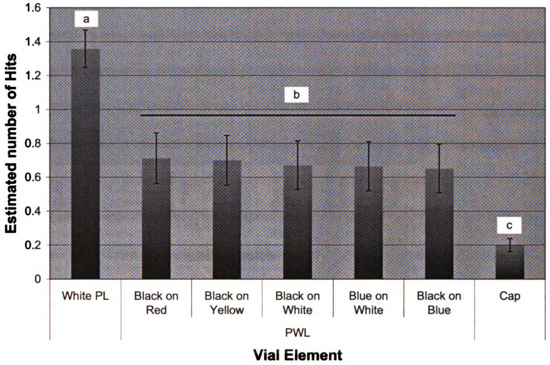


Figure 19 - Estimated number of hits on each zone ± estimated standard error. Letters (a, b, c) indicate statistically significant differences ($p < 0.05$)

As expected, longer time spent examining the vial was associated with a greater number of hits. Averaged across populations and treatments, female subjects made more visual hits than male subjects on any zone ($p = 0.0099$). Also, the number of hits averaged across zones was greater for the control population than for the test population ($p = 0.0010$). The graphs below (see Figure 20 and Figure 21) show the average number of hits that were made by subjects of both genders and both populations on the seven label zones (white pharmacy label, five color contrasts of PWL's and the Cap). They reflect the findings described above.

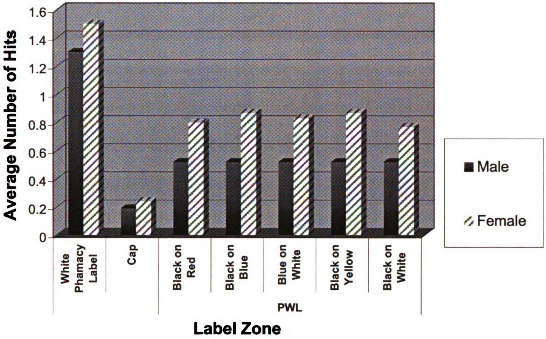


Figure 20 - Average Number of hits according to Gender

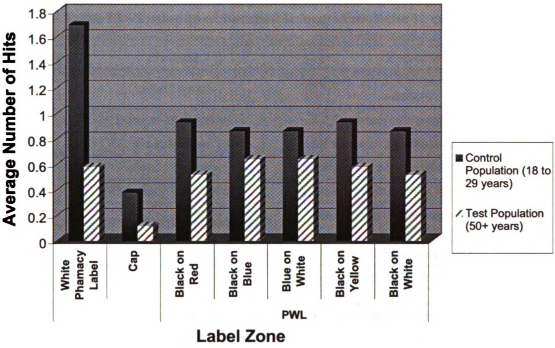


Figure 21 - Average Number of hits by Population (Control vs Test)

Recall:

Researchers encoded free form recall responses into 12 categories, namely: label layout, dosage information, drug names, patient information, pharmacy information, prescriber information, PWL color, PWL text, cap information, vial colors and medication characteristics. Recall was coded in each category for each subject in binary fashion (recall; Yes/No) and analyzed using a generalized linear mixed model. Fixed effects included in the model were: population (control vs test), response category and their 2-way interaction. Subjects within the two populations (control and test) were considered as random blocking factors in the model. Table 13 provides examples of recall information in each of the 12 categories.

Table 13 - Examples of Recalled Information in the 12 categories

Category	Element	Example
White Pharmacy Label	Label Layout	"There was some dosage information in the bottom"
	Dosage Information	"take two pills twice daily..."
	Drug Names	"I think one of the pills was Wellbutyrin"
	Patient Information	"...I remember seeing an Amy on there ...."
	Pharmacy Information	"..all vials were from the Olin health center.."
	Prescriber Information	"..One of the doctors was called Amy something.."
PWL	PWL Color	"...One of the labels was red..."
	PWL Text	"..one label said not to drink when taking the pill..."
Cap	Cap Information	"...press down and turn on all the caps.."
Other	Vial Colors	"...the vials were yellow in color.."
	Medication Characteristics	"...there were small white tablets in the vial.."

Gender, ethnicity, age, health literacy, color blindness, number of prescription drugs per day and visual acuity were initially considered for inclusion into the model. However, none of these variables showed a p-value < 0.05 when included in the model and, thus, were excluded from further analysis.

Table 14 – Predicted Probability of Recall of Information in a particular category (Data for Figure 21)

Category	Element	Probability of Recall (%)	Standard Error
White Pharmacy Label	Label Layout	51.82	10.2
	Dosage Information	68.97	8.468
	Drug Names	55.93	9.285
	Patient Information	58.32	9.622
	Pharmacy Information	55.22	9.903
	Prescriber Information	40.68	9.088
PWL	PWL Color	43.76	9.148
	PWL Text	53.18	9.2
Cap	Cap Information	21.31	7.43
Other	Vial Colors	29.94	8.867
	Medication Characteristics	10.82	5.952

Results are presented here in the original scale as 'least square mean' estimates of the probability of recall and estimated standard errors at 95% significance (see Table 14). Model fitting and parameter estimation was conducted using the GLIMMIX procedure of the statistical software SAS (version 9.1, SAS Institute, Cary, NC).

When averaged across population (control vs test), the proportion of subjects that were able to recall a given category was influenced by category itself ($p = 0.0008$). Post-hoc pair wise comparisons were performed using Tukey-Kramer's adjustment to avoid inflation of type I error rate (See Table 14 and Figure 22). Figure 22 indicates dosage information as the most frequently recalled category. A significant interaction between population (control versus test) and recall category on the probability of recall ($p = 0.21$) was not concluded. A breakdown of the raw data for this dependant variable is provided in Appendix 4.

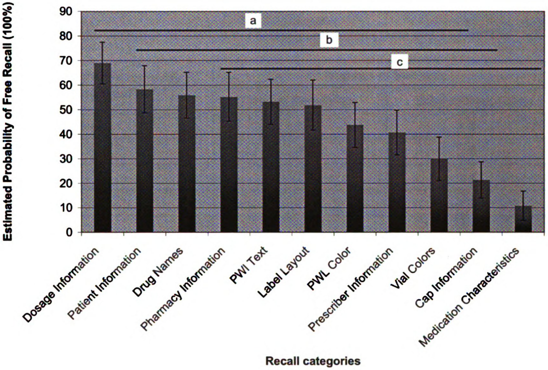


Figure 22 - Predicted probability of recall + standard error. Letters (a, b c) indicate statistically significant difference ($p < 0.05$)

Recognition:

The number of PWL's correctly recognized by the subjects were noted and counted (Table 15 and Figure 23).

It was mentioned in the materials and methods chapters that subjects were presented with a recognition sheet with ten labels (see Figure 13) and asked to pick out the five labels that they had just viewed on the prescription vials. On this recognition sheet, five PWL's were the actual PWL's that subjects had viewed in the prescription vials. The other five PWL's were labels which had the same text, but different color contrasts. The tabulation below indicates the number of PWL's that subjects recognized in both populations (Table 15 and Figure 23). A breakdown of the raw data has been provided in Appendix 5.

Table 15 - Number of PWL's correctly recognized (Data for Figure 23)

Number of PWL's correctly recognized	Number of Subjects	
	Control Population	Subject Population
None	1	1
1	1	8
2	2	4
3	7	2
4	1	2
5	2	0

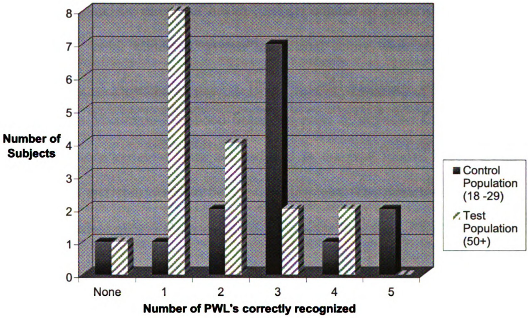


Figure 23 - Number of PWL's correctly recognized

For the purpose of statistical analysis, subjects were scored in this recognition test in two ways.

The number of PWL's out of the original five that were correctly recognized was counted and subjects were given a score out of 5. This was one dependant variable. A subject who had recognized all five PWL's correctly would have a score of five and a subject who had recognized none would have a score of 0.

A second dependant variable was created by counting both the number of PWL's correctly recognized and the number of PWL's correctly ignored. For example, a subject who recognized all five PWL's correctly, while ignoring the other five correctly received a score of 10.

For both dependant variables explained above, a general linear model was deemed to be an appropriate fit. This model included the fixed effect of population (control vs test). In addition, the effects of health literacy, number of prescription drugs per day, ethnicity, age, color blindness, gender and visual acuity were considered for model inclusion. Neither of these explanatory variables made significant contribution to model fit based on a Maximum-likelihood based Bayesian Information Criteria (BIC). Thus, these explanatory variables were not included in the final model. Model assumptions were checked and considered to be appropriately met. The model was fitted using the MIXED procedure of SAS (Version 9.1, SAS Institute Inc., Cary, NC).

Table 16 - Estimated Least Square Means and Standard Error for Recognition Scores /5 and /10

Population	Estimated Least Square Means of recognition test scores		Standard Error	
	Score / 5	Score / 10	Score / 5	Score / 10
Control (18 to 29 years)	2.66	7.13	0.34	0.45
Test (50 + years)	1.77	5.11	0.32	0.42

Table 16 indicates the estimated lease square means and standard errors for both dependant variables (/5 and /10). Pair wise comparisons were conducted using Bonferroni adjustment to avoid inflation of Type I error rate.

For the first dependant variable where subjects were scored out of five, a marginal main effect of population (control vs test) was identified on the number of correctly recognized PWL's ($p = 0.0540$). Subjects from the control population correctly recognized more PWL's than the test population (i.e. younger subjects did a better job at recognizing the PWLs).

For the second dependant variable where subjects were scored out of ten, a significant main effect of population (control vs test) was identified on the number of correctly recognized PWL's ($p = 0.0031$). The control population correctly recognized more PWL's than the test population.

Legibility Testing:

The Lockhart legibility instrument provides a measure of legibility based on a legibility index, i.e. the degrees of rotation of the polarizing filters in the legibility instrument. These readings were collected in triplicate on a total of 33 subjects (15 control; 18 test). Please note that the total number of subjects (control + test) for all other response variables is 32, whereas the total number of subjects for legibility data is 33. This is because eye tracking and subsequent recall data was not available for one subject in the test population. A breakdown of the raw legibility data is provided in Appendix 6.

Triplicates of legibility data were averaged for each subject by PWL color contrast and were then log transformed to meet model assumptions. The transformed response variable was then modeled as a function of the fixed effects of population (control vs test), label color and their 2-way interaction using a general linear mixed model. Subject, nested within population, was fitted as a random factor in the model in order to recognize the appropriate experimental unit for population (control vs test) and the blocking factor for treatment. Visual acuity was included in the model as an explanatory covariate. Age, in its continuous form, was considered for inclusion in the model but did not improve model fit. It was thus removed from the final model. Model fitting and parameter estimation was conducted using the MIXED procedure of the statistical software SAS (version 9.1, SAS Institute, Cary, NC).

Regardless of the PWL color contrast, even after adjusting for reduced visual acuity, the test population (50+) indicated a statistically greater degree of filter rotation (poorer legibility of prescription warning labels) than the control population (18-29) ($p < 0.0001$). This is consistent with previous studies performed using this method (Bix 1998; Bix 2001), and can be explained by the physiological effects of aging that in turn cause older subjects to require more light. These include: the yellowing of the cornea, vitreous humor and reduction in pupil size as people age. Visual acuity was also significantly associated with filter rotation ($p < 0.0016$); people with poorer visual acuity scores required a greater degree of rotation of the polarizing filter to read text, such that one unit increase in visual acuity resulted in a 20% cumulative increase in filter rotation. This relationship was apparent in all PWL color contrasts for both populations (control and test).

Table 17 – Least Square Means for the averaged triplicates of Legibility Index (Data for figure 24)

PWL Color Contrast	Least Square Mean Estimates of Legibility Index	Lower Confidence Limit	Upper Confidence Limit
Black on Blue	43.129	38.1082	48.8113
Black on Red	33.338	30.275	36.711
Blue on White	24.4183	22.1747	26.8888
Black on Yellow	22.4046	20.3461	24.6714
Black on White	20.3945	18.5207	22.4579

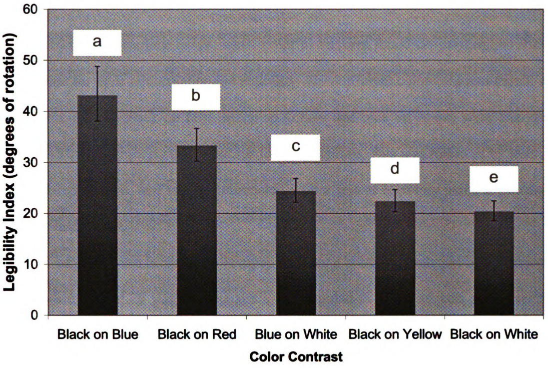


Figure 24 - Estimated legibility index and 95 % confidence interval for the five PWL color contrast combinations. Letters (a,b,c) indicate statistically significant differences)

Averaged across both populations (control and test), PWL color had a significant impact on degrees of filter rotation (P<0.0001). Results are presented in the original scale (os) as least square mean estimates and estimated 95% confidence interval. Post-hoc, pair wise comparisons were performed using Bonferroni's adjustment to avoid inflation of type I error rate (see Table 17 and Figure 24). It can be seen from Figure 24 that the black on blue color contrast is the least legible and the black on white color contrast is the most legible.

Conclusions, Limitations & Future Research

Prescription warning labels, small colorful stickers found on prescription vials, were studied for their noticeability and legibility. Five different color contrasts (black on white, black on red, black on blue, black on yellow and blue on white) were analyzed using an ASL 501 bright pupil eye tracker for noticeability. This was followed by tests of recall and recognition, and testing culminated with a test of legibility using a Lockhart Legibility Instrument.

32 subjects were broken into two populations; a control population of 15 subjects aged between 18 and 29 years of age and a test population of 17 subjects aged over 50 years. Data was obtained from these subjects and analyzed using a mixed model statistical analysis. For a complete and detailed analysis of data obtained, readers are encouraged to go through Chapter 8: Results, Analysis and Discussion.

This chapter contains an overview of the several conclusions that can be drawn from this study. The later part of this chapter contains a summary of the various limitations and possible sources of error in the study.

The conclusions below are explained in the context of future studies that are envisioned in similar topics.

Noticeability and Attentive Behavior:

Study findings do not provide evidence of color contrast promoting differences in attentive behavior on PWLs. There was not enough evidence to conclude an association between color contrast of PWL's and noticeability (see chapter 8). It is speculated that expanding the sample size and repeating this study will result in similar findings.

All the dependant variables used to quantify noticeability indicated that attentive behaviors are directed toward the large, white prescription label. It registered a higher number of hits across all populations and had a higher estimated probability of being noticed (see chapter 8). There may be features inherent to this label (For eg: large size, higher amount of information) that promote its noticeability. These must be explored further.

It is speculated that changing the position of the PWL itself may improve noticeability. A cursory visual inspection of the prescription vial reveals the presence of the white pharmacy label, regardless of the angle from which it is viewed. This is not true in the case of the narrow, vertically positioned PWL (see Figure 7). Redesigning the layout and positioning of both the labels may contribute to better noticeability of the important information contained within the PWL's. Simple tests of vial examination and recall may reveal valuable information about these new layouts.

A future study is also envisioned where warning information is integrated into the white pharmacy label in the form of small blurbs. This would enable studying of other warning features that may promote noticeability like borders and symbols as opposed to color.

The analysis of noticeability data was performed in multiple ways. At the time of publishing of this thesis, several more techniques are being studied and investigated. It is envisioned that a future study will compare the effects of data manipulation and analysis (time, time/character, time/label density, etc.) on outcomes, i.e. an analysis of the empirical methods that can be employed for labels and packaging.

The information processing model that was the basis for this study describes noticeability and legibility as two distinct and separate steps in the path to warning effectiveness. Subsequently, eye tracking and legibility testing were performed separately and different dependant variables were used to quantify the two. It is envisioned that future studies will integrate these two steps by design and analyze these quantities together. Studies shall also involve creating newer models that assess warning effectiveness.

Recall of label and vial information:

As revealed in the analyses in Chapter 8, certain elements of the label like dosage information, patient information and drug information have a statistically higher probability of being recalled when subjects have examined prescription vials. It is suggested that future studies investigate recall rates of warning information when they are integrated with dosage and patient information.

With the increase in the availability of digital printers capable of performing variable data operations, personalized printing of warning information together with patient information is also a field that is worth exploring.

Recognition of labels:

As revealed in Figure 22 in Chapter 8, data supports an association between the age and the number of labels that were correctly recognized by subjects. There must be more exploration into the management of medical information of older consumers, specifically with respect to warnings. As mentioned, therapeutic regimens are becoming increasingly complicated and increasing the cognitive loads at a time when cognitive ability is degenerating.

Legibility of PWL's:

The findings in this study are consistent with previous work that was done on legibility of printed matter. The black text on white background color contrast has had the best legibility scores for all age groups in several previous studies.

Other salient conclusions:

While analyzing eye tracking data, there was revealed to be an association of gender and number of hits per zone (see Chapter 8). This opens potential avenues of research. Gender differences on attention capture on warnings is an area that could be researched in the future. These findings also seem to coincide with those of Godfrey et al in establishing that female subjects have a greater tendency to look for warning information (Godfrey, Allender et al. 1983).

Eye tracking has several potential applications in evaluating package design and placement. Further studies are envisioned in several areas including, but not limited to anti-counterfeiting and shelf/aisle placement.

The section below highlights some of the limitations of this study.

Limitations:

Along with wearing the ASL 501 bright pupil eye tracker,subjects were made to sit at a special eye tracking table while placing their chin on a chin rest. This entire setup was not a representation of the normal position/environment in which subjects would examine prescription medications at home. This may have affected the results in some way.

The labels used in this study were totally devoid of any symbols in order to make analysis easier. The only graphical element on the entire label set was the pharmacy logo on the large white pharmacy label. Warnings in real life are almost always seen with accompanying symbols.

The information processing model that was presented in Chapter 1 indicates that a four stage process needs to be followed for a warning to be effective. It was also explained that this study would try to gain insight into noticeability and legibility of PWL's. Optimum noticeability and legibility of a warning do not necessarily imply that the warning will be easy to comprehend and comply. Further studies must investigate comprehension and compliance of PWL's in order to obtain a complete spectrum of information about these labels in the context of the information processing model.

When subjects arrived for this study, it is possible that they had in their minds, information about the labels on the prescription vials in their home. Delving into concepts like prospective (long term) and working (short term) memory, it is possible that some of the recall and recognition responses were influenced by these reserves of information. Drawing parallels to Morrow et al's work on medication information schemata (Morrow, Leirer et al. 1996), it is possible that some recall responses may be been predetermined by these schema and not actually associated with the vials that were viewed during the eye tracking study.

APPENDIX 1: CONSENT FORM

Michigan State University

School of Packaging

<u>INSTRUCTIONS AND RESEARCH CONSENT FORM – Prescription vial use experiment</u>

You are being asked to participate in a research study. The entire study should last no more than 2 hours. In exchange for your participation, you will receive $30. You may discontinue participation at any time and still keep the $30.

As part of this research, we will record your gender, ethnicity, educational background, age and the number of prescription and OTC medications that you take each day. This information will be tied to a subject number; you will not be identified by name and your privacy will be maintained to the maximum extent of the law.

We will also test your visual acuity (20/20, 20/30, 20/40, etc.) and color blindness. These tests will be conducted by asking you to view a series of cards and asking you to decipher images to the best of your ability. Following these tests, you will be then shown a card with 11 commonly used medical terms. You will be asked to say these words. This will give us an idea of your familiarity with some terms commonly used in the field of healthcare.

Instrument 1 ASL Eyetracker

After these tests have been run, we will begin the eye tracking portion of the study. The eye tracker is a very sensitive eye movement monitor, which can tell us exactly where your eyes are looking while you are viewing a package. A beam of light that cannot be detected by the human eye will be shone into your eye. The instrument tracks the movement of your eye by tracking the movement of the beam.

You will wear a light (4 ounce) camera on your head (attached to a headband). Once the eye tracking equipment is adjusted, we will prepare the eye movement equipment. You will be asked to sit at a desk. For a period of about 5 minutes you will be asked to sit as still as possible with your chin on the chinrest, and move nothing but your eyes. While holding your head and body as still as possible, you will be asked to look at certain locations in space so that the researcher can calibrate your eye's position. After we have calibrated the equipment, you may take your chin off of the chin rest and move until testing begins.

During the actual experiment you will be asked to rest your chin in the chin rest, while you examine 5 different prescription packages. You will be handed a series of five bags, one at a time, and asked to remove the contents (a prescription drug vial), and do what you normally would do if you were viewing a new medication. If you read anything on the vial, please press the vial against the glass while doing so. At the end of this test you will be asked to answer a short questionnaire based on the vials that you just viewed.

If you have any questions at any time please ask.

Instrument 2 Lockhart Legibility

After we have tested you for color blindness, measured your visual acuity and had you view a series of prescription vials using the eye tracker, you will be asked to read several labels using an instrument called the Lockhart Legibility Instrument (LLI). The LLI is a tool that provides a numerical value for how easy or difficult a message is to read. The purpose of this research is to investigate the legibility of a variety of package labels. It is our hope that this will ensure that the label information provided on a variety of labels is sufficiently legible.

You will read a label placed inside the grey box. Look into the box through the screen on the front. As you look through the screen turn the knob on the right side of the box until you can easily read the words on the card in the box without straining your eyes. The operator will record the value you get for each card. Once the value is recorded, turn the knob back to its starting position so that the screen is dark again. The operator will put a different card in the box for you to read.

Risks and Benefits

There is little or no risk associated with this research. It is possible that you will experience some discomfort while you are asked to remain still with your head on the chin rest, or when the researcher adjusts the head piece. Please let us know if you do experience any discomfort at any time. There is no direct benefit to you for participating in this study. However, we are hopeful that we can learn what constitutes well designed prescription labels so that this information is used to create future label designs that are more effective.

Confidentiality

Study results will be treated in strict confidence. Your privacy will be protected to the maximum extent allowable by law. Within these restrictions, results of this study will be made available to you at your request. Data will be stored on a password protected computer in the Packaging Building and in a locked file cabinet in Laura Bix's office for a minimum of 3 years and a maximum of 10 years. Collected data will only be stored subject number and cannot be tied to you identity.

If you have any concerns or questions about this research study, such as scientific issues, how to do any part of it, or if you believe you have been harmed because of the research, please contact the researcher Laura Bix 517-355-4556; 153 Packaging Building East Lansing MI 48824 bixlaura@msu.edu.

If you have questions or concerns about your role and rights as a research participant, would like to obtain information or offer input, or would like to register a complaint about this study, you may contact, anonymously if you wish, the Michigan State University's Human Research Protection Program at 517-355-2180, Fax 517-432-4503, or e-mail irb@msu.edu or regular mail at 202 Olds Hall, MSU, East Lansing, MI 48824.
I voluntarily agree to participate in the Eye Tracking and label legibility study.

____

Date: ______________________________

You will be provided with a copy of your signed consent form.

APPENDIX 2: RECRUITMENT ADVERTISEMENT FOR TEST POPULATION

Prescription Vial Use Experiment

Participants wanted for research concerning prescription vial use. In exchange for your participation, you will receive $30. The study will take no longer than 2 hours. To participate in this study:

- You must be able to provide your own transportation to the School of Packaging at Michigan State University
- You must be over 50 years of age
- You must not be legally blind
- Must not wear hard contact lenses
- Be willing to provide a contact phone number or email so that researchers can contact you to remind you of your appointment

You are being asked to participate in a study of prescription drug use being conducted by graduate student Raghav Prashant Sundar. As part of his Master's Program, Raghav is investigating the use of prescription vials. He is using two instruments to do this. Instrument 1 is an ASL eye tracker. If you choose to participate, you will be hooked up to the eye tracker which will track your eye movements as you look at 5 prescription vials. You will also be asked to answer a brief questionnaire about the prescription drug vials you just viewed. Instrument set up and the test itself should take no longer than 45 minutes.

Prior to testing, your visual acuity (20/20,20/30, 20/40, etc.) and a test for color blindness will be administered. These tests involve viewing a series of cards. You will also be shown a list of eleven health terms and asked to read these terms aloud. This will give us an idea of your familiarity with some terms commonly used in the field of healthcare. We will also ask you to fill out a brief survey which includes information about your race, gender, age, educational background and the number of prescription drugs and OTC drugs that you consume per week.

The final portion of the research involves a test of the legibility of the labels that you viewed on the vials previously. This portion of the study is done on the "Lockhart Legibility Instrument" (LLI), an instrument that was developed at the School of Packaging. For this portion of the assignment, you will be asked to rotate a small wheel located to your right as you are seated in front of the instrument until the first point that you can easily read the words on the label. This portion of the test should take no longer than 25 minutes.

If at any time you are uncomfortable with the testing or wish to discontinue the data collection process, you may discontinue participation without penalty.

If you are interested in pursuing this opportunity, please contact Raghav Prashant Sundar at sundarra@msu.edu or cell phone 517-898-9029 to make an appointment.

If you have questions or comments regarding this study, please contact Dr. Laura Bix, Assistant Professor or Packaging at Michigan State University at 517-355-4556 or bixlaura@msu.edu.

APPENDIX 3: DATA RECORDING SHEET

Human Factors & Medical Packaging Research Group, School of Packaging
Research Study (IRB **IRB# 08-246/ APP# i029464)**

Level 1: Personal Information

Subject Id: ____________________ Age: ____________________

Ethnicity: ____________________ Highest level of education: ____________________

Number of Prescription Drugs per week: ______________________

Level 2: Eye Tracking <Please turn over for Recall Test forms(free and guided)>

Level 3: Pre-Testing

Literacy Score:

Hold document at convenient reading distance and say. "I want to hear you read as many words as you can from this list. Begin with the first word in List 1 and read aloud. When you come to a word you cannot read, do the best you can or say, 'blank' and go onto the next word." (6 or less has poor health literacy)

Fat Flu Pill Allergic Jaundice Anemia Fatigue

Directed Colitis Constipation Osteoporosis

Visual Acuity:

I want you to hold this card at about 16 inches from your eyes and try to read the lowest line on this card.

20/100:	**S D K H N**	**O C V R Z**
20/150:	**H C Z S V**	**D O R N K**
20/100:	**O N H D R**	**Z K S C V**
20/80:	**Z V R C K**	**N O H D S**
20/60:	**K N C S O**	**V H Z R D**
20/50:	**R V D K C**	**S H O N Z**
20/40:	**H N D C R**	**V S K Z O**
20/30:	**Z C O N S**	**D H R V K**
20/20:	**C K D Z H**	**R N O S V**

Result: 20/____

Color Blindness Test:

6	**42**	**56**	**57**	**75**
5	**3**	**56**	**27**	**89**
86	**15**	**74**	**47**	

10 and above means "normal color vision"

Level 4: Legibility Instrument

Label	Trial 1	Trial 2	Trial 3
Dummy 1			
Dummy 2			

Level 5: Compensation

Appendix 4: Raw Data for Recall Test

Table 18 - Information recalled by subjects of the control population in 12 categories in binary form (recalled/not recalled).

Subject	Vial Colors	Cap Information	Number of Vials	Drug Names	Prescriber's Information	Patient Information	Pharmacy Name	Dosage Information	Warning Text	Warning Label Colors	Label Layout	Medication Characteristics
1	0	0	0	1	1	0	1	1	0	0	1	0
2	1	0	0	1	1	1	1	1	1	1	1	0
3	0	0	0	0	1	1	1	1	1	1	1	1
4	0	0	0	0	0	0	0	1	0	0	0	0
5	1	0	0	1	0	1	0	1	0	0	1	0
6	1	1	0	1	1	1	1	1	0	0	1	0
7	1	1	0	0	0	0	1	1	1	0	0	0
8	0	0	0	1	0	1	1	0	1	1	0	0
9	1	0	0	1	0	1	0	0	0	1	1	0
10	0	0	0	0	1	1	1	1	0	0	1	1
11	0	0	0	1	0	1	1	0	1	1	1	0
12	1	0	0	0	0	0	0	0	0	0	1	0
13	1	0	0	0	1	1	1	1	1	0	1	0
14	0	0	0	0	1	1	1	1	1	1	0	0
15	0	1	0	0	0	1	1	0	1	1	1	1

Table 19 - Information recalled by subjects of the test population in 12 categories in binary form (recalled/not recalled).

Subject	Vial Colors	Cap Information	Number of Vials	Drug Names	Prescriber's Information	Patient Information	Pharmacy Name	Dosage Information	Warning Text	Warning Label Colors	Label Layout	Medication Characteristics
16	0	0	0	1	0	1	1	1	1	0	0	0
18	1	0	0	0	1	0	0	0	0	1	1	0
19	0	0	0	1	0	0	0	1	1	0	0	0
20	0	1	0	1	0	0	1	1	1	1	0	0
21	1	1	0	0	0	0	1	1	1	1	1	0
22	0	0	0	0	0	0	0	1	0	1	1	0
23	0	0	0	0	0	0	0	1	1	0	0	0
24	0	0	0	1	0	0	0	1	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	1	0
26	0	0	0	1	0	1	0	1	1	1	0	0
27	0	0	0	1	1	1	0	1	0	0	0	0
28	1	1	0	0	1	1	0	0	1	1	0	0
29	0	1	0	1	0	0	0	1	1	0	0	1
30	0	0	0	1	1	1	1	1	0	0	0	0
31	0	0	0	1	1	1	1	1	0	0	0	0
32	0	0	0	1	1	1	1	0	0	0	1	0
33	0	0	0	1	0	0	0	0	1	1	0	0

Appendix 5 – Raw Data for Recognition Test

Table 20 - Recognition Data for subjects of the control population for all 10 labels in binary form (circled/did not circle)

Subject	Original PWL's on vials examined by subjects					Extra PWL's on recognition sheet					Recognition Dependant Variables	
	Black on Red	Black on Yellow	Black on Blue	Blue on White	Black on White	Black on Purple	Black on Orange	Black on Green	Black on Pink	Black on Beige	Total/5	Total/10
1	0	1	0	0	1	0	0	0	0	0	2	7
2	1	1	1	1	1	0	0	0	0	0	5	10
3	0	1	0	1	1	0	0	0	0	0	3	8
4	1	1	0	0	1	0	0	0	0	0	3	8
5	0	0	0	0	0	0	0	0	1	0	0	4
6	0	0	1	0	1	0	0	0	1	0	2	6
7	1	0	1	0	1	1	0	0	1	0	3	6
8	1	0	0	0	0	0	0	0	0	0	1	6
9	1	1	0	0	1	1	0	1	0	0	3	6
10	1	1	1	0	1	1	0	0	0	0	4	8
11	1	1	1	1	1	0	0	0	0	0	5	10
12	0	0	0	0	0	0	0	0	0	0	0	10
13	1	1	0	1	0	1	1	0	0	0	3	6
14	0	1	1	0	1	1	1	0	0	0	3	6
15	0	1	1	0	1	0	1	0	1	0	3	6

Table 21 - Recognition Data for subjects of the test population for all 10 labels on the recognition sheet in binary form (circled/did not circle)

Subject	Original PWL's on vials examined by subjects					Extra PWL's on recognition sheet					Recognition Dependant Variables	
	Black on Red	Black on Yellow	Black on Blue	Blue on White	Black on White	Black on Purple	Black on Orange	Black on Green	Black on Pink	Black on Beige	Total/5	Total/10
16	0	1	1	1	1	0	0	0	1	0	4	8
18	0	1	0	0	0	0	0	0	0	0	1	6
19	1	0	1	0	0	1	0	1	0	0	2	5
20	1	0	0	0	0	1	0	1	1	1	1	2
21	0	1	0	0	0	1	1	1	0	1	1	2
22	0	1	0	0	0	0	0	0	1	0	1	5
23	0	0	0	0	1	0	1	0	1	0	1	4
24	1	0	0	0	1	0	0	0	0	0	2	7
25	0	1	1	1	1	0	0	0	0	1	4	4
26	1	1	0	1	0	0	0	0	0	0	3	8
27	0	0	0	0	0	0	0	0	1	0	0	4
28	1	0	0	0	1	1	0	1	0	0	2	5
29	0	0	1	1	0	0	1	0	1	1	2	4
30	0	0	0	0	0	0	0	0	0	0	0	5
31	0	0	0	0	1	0	0	0	0	0	1	6
32	0	1	0	0	0	0	0	0	0	0	1	6
33	1	0	1	0	1	1	1	0	0	0	3	6

Appendix 6: Average Legibility Indices

Table 22 - Average of Legibility Index triplicates for all subjects. Subjects 1-15 belong to the control population. Subjects 16 to 33 belong to the test population.

Subject	Age	PWL color contrast				
		Black on Red	Black on Blue	Blue on White	Black on Yellow	Black on White
1	21	26.7	34.6	19.3	16.2	13.96
2	20	15.96	20.86	11.93	10.53	10.23
3	22	18.6	24.26	13.13	11.83	10.96
4	23	28.13	34.9	21.63	17.66	18.33
5	21	16.2	21.33	12.56	12.96	11.4
6	21	25.13	30	16.13	16.66	16.6
7	22	24	38.03	16.86	14.93	16.03
8	24	26.83	35.73	20.13	18.66	17.46
9	22	15.86	15.43	11.43	10.26	9.43
10	22	22.73	27.9	15.33	12.7	12.76
11	24	22.36	29.3	17.23	14.5	13
12	21	18.7	15.96	14.8	14.3	11.93
13	26	25.03	45.5	20.73	16.06	15.53
14	24	26.6	41.63	19.26	18.46	15.4
15	29	15.46	18.76	11.8	10.86	10.3
16	70	52.33	62.5	32.66	25.53	21.26
17	71	63.23	83.33	44.46	47	38.16
18	70	90	90	64.83	56.8	55.6
19	53	34.93	89.4	31.96	31.16	24.9
20	54	34.5	48.23	23.46	21.13	18.9
21	64	90	90	52.56	55.8	45.9
22	58	35.73	87	30.53	27.6	25.1
23	51	48.3	44.5	36.76	35.86	31.63
24	77	90	90	90	90	90
25	74	33.56	38.26	23.2	20.6	20.46
26	51	63.9	90	45.8	46.36	37.73
27	71	33.03	39.06	22.46	22.7	16.53
28	63	49.03	87.4	28.13	27.66	26.56
29	61	60.5	82.36	39.2	37.66	31
30	55	51.66	53.13	38.86	32.93	31.66
31	64	34.83	49.9	27.23	26.03	23.23
32	64	67.43	90	57	52.1	46.1
33	59	41.33	43.33	28.66	24	24

BIBLIOGRAPHY

Ahlbaum, A. (2008). "Eye Tracking breaks into shopper research with major Tobii order." Retrieved Feb 4th, 2009, from http://www.tobii.com/all_news/eye_tracking_breaks_into_shopper_research_with_major_tobii_order.aspx?profile=17741&menu=17757&submenu=17888.

Ajzen, I. and M. Fishbein (1977). "Attitude-Behavior Relations - Theoretical-Analysis and Review of Empirical-Research." Psychological Bulletin **84**(5): 888-918.

Anders, G. (2001). PIlot's Attention allocation during approach and landing: Eye and Head-tracking research in an A330 full flight simulator. International symposium on aviation psychology (ISAP). Columbus, OH.

Andrews, J. C. (1995). "The Effectiveness of Alcohol Warning Labels - a Review and Extension." American Behavioral Scientist **38**(4): 622-632.

Asaad, W. F., G. Rainer, et al. (2000). "Task-specific neural activity in the primate prefrontal cortex." Journal of Neurophysiology **84**(1): 451-459.

Ascione, F. J. and L. A. Shimp (1984). "The Effectiveness of 4 Education Strategies in the Elderly." Drug Intelligence & Clinical Pharmacy **18**(11): 926-931.

Aspden, P. and Institute of Medicine (U.S.). Committee on Identifying and Preventing Medication Errors. (2007). Preventing medication errors. Washington, DC, National Academies Press.

ASTM. (2006). "Standard Test Method for Measurement of Comparative Legibility by Means of Polarizing Filter Instrumentation." from http://www.astm.org/Standards/D7298.htm.

Ault, H. (2007). "A warning about prescription drug labels." Retrieved 01/13/09, 2009.

Baber, C. and J. Wankling (1992). "An Experimental Comparison of Text and Symbols for in-Car Reconfigurable Displays." Applied Ergonomics **23**(4): 255-262.

Backinger, C. L. and P. A. Kingsley (1993). Write it right: Recommendations for developing user instruction manuals for medical devices used in home health care. Washington, DC, US Department of Health and Human Services.

Bass, P. F., J. F. Wilson, et al. (2003). "A shortened instrument for literacy screening." Journal of General Internal Medicine **18**(12): 1036-1038.

Bates, D. W. (2007). "Preventing medication errors: A summary." American Journal of Health-System Pharmacy **64**: S3-S9.

Berman, A. (2004). "Reducing medication errors through naming labelling and packaging." Journal of Medical Symptoms **28**(1): 9-29.

Bix, L. L. (1998). The Effect of subject on legibility. School of Packaging. East Lansing, Michigan State University. **Masters**.

Bix, L. L. (2001). Towards a performance standard for typeface legibility: The Lockhart Legibility Instrument. School of Packaging. East Lansing, Michigan State University. **PhD**.

Bix, L. L. (2003). "The effect of color contrast on message legibility." Journal of Design Communication(5).

Boyd, S. T., L. C. Boyd, et al. (2006). "Medication therapy management survey of the prescription drug plans." Journal of the American Pharmacists Association **46**(6): 692-699.

Brown, C. S., B. L. Solovitz, et al. (1988). "Short-Term and Long-Term Effects of Auxiliary Labels on Patient Knowledge of Precautionary Drug Information." Drug Intelligence & Clinical Pharmacy **22**(6): 470-474.

Bureau, U. C. (1997). Aging in the United States: Past, Present and Future.

Cahill, M. C. (1976). "Design-Features of Graphic Symbols Varying in Interpretability." Perceptual and Motor Skills **42**(2): 647-653.

Chapman, P. R. and G. Underwood (1998). Visual search in dynamic scenes: Event types and the role of experience in viewing driving situations. Eye Guidance in Reading and Scene perception. G. Underwood. Oxford, England, Elsevier: 369-394.

Charness, N. and E. A. Bosman (1992). Human Factors and Age. Handbook of aging and cognition. T. A. Salthouse and F. I. M. Craik. Hillsdale, NJ, Lawrence Ernbaum Associates: 495-552.

Crespo, A., R. Cabestrero, et al. (2007). "Visual attention to health warnings in tobacco advertisements: An eye-tracking research between smokers and non-smokers." Studia Psychologica **49**(1): 39-51.

d'Yedwalle, G., G. Desmet, et al. (1998). Film perception: The processing of film cuts. Eye Guidance in reading and scene perception:. G. Underwood. Oxford, England, Elsevier: 357-368.

Davis, T. C., R. Michielutte, et al. (1998). "Practical assessment of adult literacy in health care." Health Education & Behavior **25**(5): 613-624.

Davis, T. C., M. S. Wolf, et al. (2006). "Literacy and misunderstanding prescription drug labels." Annals of Internal Medicine **145**(12): 887-894.

De La Fuente, J. (2006). The use of a universal design methodology for developing child-resistant drug packaging. School of Packaging. East Lansing, Michigan State University. **Masters in Packaging**.

DeJoy, D. M. (1991). "A Revised Model of the Warnings Process Derived from Value-Expectancy Theory." Proceedings of the Human Factors Society 35th Annual Meeting: 1043-1047.

DeJoy, D. M. (1999). Attitudes and Beliefs. Warnings and Risk Communication. M. S. Wogalter, D. M. DeJoy and K. R. Laughery. London ; Philadelphia, Taylor & Francis: 189 - 213.

DeJoy, D. M. and K. R. Laughery (1999). Preface. Warnings and Risk Communication. M. S. Wogalter, K. R. Laughery and D. M. DeJoy. Philadelphia, PA, Taylor and Francis: xiii-xvi.

Dockeray, F. C. (1910). "The span of vision in reading and the legibility of letters." Journal of Educational Psychology **1**: 123-131.

Dombrowski, R. C., A. A. Schuna, et al. (1991). "Auxiliary label use in pharmacists in South Carolina." Journal of Pharmacy Technology **7**: 64 - 70.

Dubay, W. H. (2004). Principles of Readability. Costa Mesa, CA.

Duchowski, A. T. (2003). Eye tracking methodology : theory and practice. New York, Springer.

Edwards, M. L. and N. C. Ellis (1976). "Evaluation of Texas Driver Improvement Training-Program." Human Factors **18**(4): 327-333.

Ells, J. G., R. E. Dewar, et al. (1980). "An evaluation of six configurations of the railway crossbuck sign." Ergonomics **23**: 359-367.

FDA (1999). Final rule - Over the Counter Human Drugs Labelling requirement.

FDA (2004). Food Allergen Labeling and Consumer Protection Act. USA.

Ferguson, J. M., R. Discenza, et al. (1987). "Increasing the odds of patient compliance through prescription warning labels." Journal of health care marketing **7**(1): 37-46.

Fischer, P. M., J. W. Richards, et al. (1989). "Recall and Eye Tracking Study of Adolescents Viewing Tobacco Advertisements." Jama-Journal of the American Medical Association **261**(1): 84-89.

Fischhoff, B. (1977). "Cognitive Liabilities and Product Liability." Journal of Products Liability **1**(3): 207-219.

Flesch, R. (1948). "A new readability yardstick." Journal of Applied Psychology **32**: 221-233.

Franklin, D. (2005). The Consumer: And Now, a Warning about Labels. The New York Times. New Yor.

Gamlin, P. and D. Tweig. (1996). "CRI: A Combined Visual Display and Eye Tracking System for High-Field fMRI Studies " Retrieved February 4th, 2009, from http://www.nsf.gov/awardsearch/showAward.do?AwardNumber=9634422.

Godfrey, S. S., L. Allender, et al. (1983). Warning Messages: Will the consumer bother to look? Human Factors Society 27th Annual Meeting, Santa Monica, CA.

Godfrey, S. S. and K. R. Laughery (1984). "The biasing effects of product familiarity on consumers' awareness of hazard." Proceedings of the Human Factors Society 28th Annual Meeting: 388-392.

Goldhaber, G. M. and M. A. Deturck (1989). "A developmental analysis of warning signs: the case of familiarity and gender." Proceedings of the Human Factors Society 33rd Annual Meeting.

Gryfebecker, B. M., H. J. Segal, et al. (1989). "Effect of Auxiliary Prescription Labels on the Elderly Ambulatory Patients Drug Knowledge." Dicp-the Annals of Pharmacotherapy **23**(4): 324-329.

Harris, J. L., W. A. Rogers, et al. (1998). "Written language comprehension in younger and older adults." Journal of Speech Language and Hearing Research **41**(3): 603-617.

Hellmich, N. (2008). Aging population making more visits to the doctor's. USA Today.

Hill, R. D. (1993). "Predictors of Reading Performance in College-Educated Older Adults." Educational Gerontology **19**(1): 21-30.

Hunn, B. P. and T. A. Dingus (1992). "Interactivity, Information, and Compliance Cost in a Consumer Product Warning Scenario." Accident Analysis and Prevention **24**(5): 497-505.

Janynes, L. S. and D. B. Boles (1990). A behavioral study of warning labels for consumer products: Perceived danger and use of pictographs. Human Factors and Ergonomics Society 34th Annual Meeting, Santa Monica, CA, Human Factors Society.

Johnston, M., A. Clarke, et al. (1986). "Facilitating Comprehension of Discharge Medication in Elderly Patients." Age and Ageing **15**(5): 304-306.

Kalsher, M. J., M. S. Wogalter, et al. (1996). "Pharmaceutical container labels: Enhancing preference perceptions with alternative designs and pictorials." International Journal of Industrial Ergonomics **18**(1): 83-90.

Karwowski, W., J. W. Yates, et al. (1991). Advances in industrial ergonomics and safety 3 : proceedings of the Annual International Industrial Ergonomics and Safety Conference held in Lake Tahoe, Nevada, 10-14 June 1991 : the official conference of the International Foundation for Industrial Ergonomics and Safety Research. London ;, Taylor & Francis.

Kaufman, D. W., J. P. Kelly, et al. (2006). "Patterns of medication use in US adults and children, 2005: A report from the slone survey." Pharmacoepidemiology and Drug Safety **15**: S65-S66.

Kemper, S. (1992). Language and Aging. Handbook of aging and cognition. T. A. Salthouse and F. I. M. Craik. Hillsdale, NJ, Lawrence Ernbaum Associates: 495-552.

Kline, D. W. and C. T. Scialfa (1997). Sensory and Perceptual functioning: Basic research and human factors implications. Handbook of human factors and the older adult. A. D. Fisk and W. A. Rogers. New York, Academic Press: 27-54.

Krugman, D. M., R. J. Fox, et al. (1999). "Do cigarette warnings warn? Understanding what it will rake to develop more effective warnings." Journal of Health Communication **4**(2): 95-104.

Krugman, D. M., R. J. Fox, et al. (1994). "Do Adolescents Attend to Warnings in Cigarette Advertising - an Eye-Tracking Approach." Journal of Advertising Research **34**(6): 39-52.

Laughery, K. R. and A. Hammond (1999). Overview. Warnings and Risk Communication. M. S. Wogalter, D. M. DeJoy and K. R. Laughery. London ; Philadelphia, Taylor & Francis: 3 - 12.

Lehto, M. and J. Miller (1988). "Effectiveness of Warning Labels." Journal of Products Liability **11**: 225-270.

Lohse, G. L. (1997). "Consumer eye movement patterns on yellow pages advertising." Journal of Advertising **26**(1): 61-73.

Luckiesh, M. and F. K. Moss (1938). "Visibility and Readability of print on white and tinted papers." Sight-Saving Review **8**: 123-134.

Luckiesh, M. and F. K. Moss (1939). "Visibility and Readability of printed matter." Journal of Applied Psychology **23**: 645-659.

Madden, S. (1999). The law related to warnings. Warnings and Risk Communication. M. S. Wogalter, D. M. DeJoy and K. R. Laughery. London ; Philadelphia, Taylor & Francis: 3 - 12.

Mazzullo, J. M., L. Lasagna, et al. (1974). "Variations in Interpretation of Prescription Instructions - Need for Improved Prescribing Habits." Jama-Journal of the American Medical Association **227**(8): 929-931.

McLean, R. (1980). The Thames and Hudson manual of typography. London, Thames and Hudson.

Meyer, B. J. F., C. J. Young, et al. (1989). Memory improved : reading and memory enhancement across the life span through strategic text structures. Hillsdale, N.J., Lawrence Erlbaum Associates.

Miller, J. A. (1978). Labeling research : the state of the art. Cambridge, Mass., Marketing Science Institute.

Morrell, R. W., D. C. Park, et al. (1989). "Quality of Instructions on Prescription Drug Labels - Effects on Memory and Comprehension in Young and Old Adults." Gerontologist **29**(3): 345-354.

Morrow, D. G., V. O. Leirer, et al. (1996). "Medication instruction design: Younger and older adult schemas for taking medication." Human Factors **38**(4): 556-573.

NIH (2006). Dramatic Changes in U.S. Aging Highlighted in New Census, NIH Report, National Institutes of Health.

Otsubo, S. M. (1988). "A behavioral study of warning labels for consumer products: perceived danger and use of pictographs." Proceedings of the Human Factors Society 32nd Annual Meeting: 536 - 540.

Paine, F. A. and Institute of Packaging. (1963). Fundamentals of packaging. London, Blackie.

Pharmex. (2009). "Top 40 Pharmex English and Spanish Warning Labels." Retrieved 03/21/2009, 2009, from http://www.pharmex.com/PharmacyWarningLabels/WarningLabels/wleng.asp.

Pietrowski, J. (1993). Development of a methodology to quantify package and label legibility. School of Packaging. East Lansing, Michigan State University. **Masters**.

Pyke, R. L. (1926). Report on the legibility of print. London,, H. M. Stationery Off. [printed by Johnson, at the University Press.

Rayner, K., Ed. (1992). Eye movements and visual cognition: Scene perception and Reading. New York, NY, Springer-Verlag.

Rehder, T. L., L. K. Mccoy, et al. (1980). "Improving Medication Compliance by Counseling and Special Prescription Container." American Journal of Hospital Pharmacy **37**(3): 379-385.

Report (2003). Public Health and Aging: Trends in Aging --- United States and Worldwide, Morbidity and Mortality Weekly Report.

Rogers, R. W. (1983). Cognitive and physiological processes in fear appeals and attitude change: A revised theory of protection motivation. Social psychophysiology. J. Cacioppo and R. Petty. New York, Guilford. **153-176**.

Rogers, W. A., G. Rousseau, et al. (1999). Maximizing the effectiveness of the warning process: Understanding the variables that interact with age. Processing of medical information in aging patients : cognitive and human factors perspectives. D. C. Park, R. W. Morrell and K. Shifren. Mahwah, NJ, Lawrence Erlbaum: ix, 332 p.

Rosbergen, E., M. Wedel, et al. (1990). Analysing visual attention to repeated print advertising using scanpath theory, University Library Groningen, SOM Research School.

Salthouse, T. A. (1991). Theoretical perspectives on cognitive aging. Hillsdale, N.J., L. Erlbaum Associates.

Sanford, E. C. (1888). "The Relative legibility of small letters." American Journal of Psychology **1**: 402-435.

Shannon, C. E. (1948). "A Mathematical Theory of Communication." Bell System Technical Journal **27**(3): 379-423.

Sheahan, N. F., D. Coakley, et al. (1993). "Ocular Microtremor Measurement System - Design and Performance." Medical & Biological Engineering & Computing **31**(3): 205-212.

Silver, N. C. and M. S. Wogalter (1990). "Strength and Understanding of signal words in elementary and middle school students." Proceedings of the Human Factors Society 35th Annual Meeting: 580-583.

Snodderly, D. M., I. Kagan, et al. (2001). "Selective activation of visual cortex neurons by fixational eye movements: Implications for neural coding." Visual Neuroscience **18**(2): 259-277.

Soroka, W. and Institute of Packaging Professionals. (1995). Fundamentals of packaging technology. Herndon, Virginia, Institute of Packaging Professionals.

Sumner, F. C. (1932). "Influence of color on legibility of copy." Journal of Applied Psychology **16**: 201-204.

Swindell, J. A. (1999). Measuring visual search reaction time and accuracy for a product warning label function of icon, color, column and vertical placement. Raleigh, North Carolina State University.

Tinker, M. A. (1946). "Validity of Frequency of Blinking as a Criterion of Readability." Journal of Experimental Psychology **36**(5): 453-460.

Tinker, M. A. (1963). Legibility of print. Ames,, Iowa State University Press.

United States Court of Appeals, F. C. (1974). "Terri Lee DREISONSTOK, an infant, by her mother and next friend, Catherine A. Dreisonstok, and Catherine A.Dreisonstok, Appellees, (Vs). VOLKSWAGENWERK, A.G., a/k/a Volkswagenwerk Aktiegesellschaft, a foreign corporation, and Volkswagen of America, Inc., a New York corporation, Appellants." Retrieved January 9th, 2009, from http://bulk.resource.org/courts.gov/c/F2/489/489.F2d.1066.73-1074.html.

Weiss, A. P. (1917). "The focal variator." Journal of Experimental Psychology **2**: 106-113.

Wiederholt, J. B., J. A. Kotzan, et al. (1983). "The Effectiveness of Auxiliary Prescription Labels - a Pilot-Study." Drug Intelligence & Clinical Pharmacy **17**(3): 216-221.

Williams, M. V., R. M. Parker, et al. (1995). "Inadequate Functional Health Literacy among Patients at 2 Public Hospitals." Jama-Journal of the American Medical Association **274**(21): 1677-1682.

Wilson, J. and L. Hogan (1983). "Readability Testing of Auxiliary Labels." Drug Intelligence & Clinical Pharmacy **17**(1): 54-55.

Wogalter, M. S., D. M. DeJoy, et al. (1999). Organizing Theoritical Framework. Warnings and Risk Communication. M. S. Wogalter, D. M. DeJoy and K. R. Laughery. London ; Philadelphia, Taylor & Francis: 15-23.

Wogalter, M. S., D. M. DeJoy, et al. (1999). Preface. Warnings and Risk Communication. M. S. Wogalter, D. M. DeJoy and K. R. Laughery. London ; Philadelphia, Taylor & Francis: xiii.

Wogalter, M. S. and T. A. Dingus (1999). Methodology for evaluating behavioral intentions. Warnings and Risk Communication. M. S. Wogalter, D. M. DeJoy and K. R. Laughery. London ; Philadelphia, Taylor & Francis: 53-77.

Wogalter, M. S., G. A. Fontenelle, et al. (1985). Behavioral effectiveness of warnings. Human Factors Society 29th Annual Meeting, Santa Monica, CA.

Wogalter, M. S., S. S. Godfrey, et al. (1987). "Effectiveness of Warnings." Human Factors **29**(5): 599-612.

Wogalter, M. S., S. W. Jarrard, et al. (1994). "Influence of Warning Label Signal Words on Perceived Hazard Level." Human Factors **36**(3): 547-556.

Wogalter, M. S. and K. R. Laughery (1996). "Warning! Sign and Label Effectiveness." Current Directions in Psychological Science **5**: 33-37.

Wong, I. C. K. (1999). "Readability of patient information leaflets on antiepileptic drugs in the UK." Seizure-European Journal of Epilepsy **8**(1): 35-37.

Wooding, D. (2002). Fixation Maps: Quantifying Eye-Movement Traces. Eye Tracking Research and Applications (ETRA). New Orleans, LA.

Wright, P., P. Creighton, et al. (1982). "Some Factors Determining When Instructions Will Be Read." Ergonomics **25**(3): 225-237.

Yarbus, A. L. (1967). Eye movements and vision. New York,, Plenum Press.

Young, S. L. and D. Lovoll (1999). Intermediate processing stages: Methodological Considerations for Research on Warnings. Warnings and Risk Communication. M. S. Wogalter, D. M. DeJoy and K. R. Laughery. London ; Philadelphia, Taylor & Francis: xiii.

Young, S. L. and M. S. Wogalter (1990). Judgements of Hazard, Risk and Danger: Do they Differ? Human Factors and Ergonomics Society 34th Annual Meeting, Santa Monica, CA, Human Factors Society.

Zuccollo, G. and H. Liddell (1985). "The Elderly and the Medication Label - Doing It Better." Age and Ageing **14**(6): 371-376.