This is to certify that the

thesis entitled

The Effect of Subject Age on Legibility

presented by

Laura Bix

has been accepted towards fulfillment
of the requirements for

Master's degree in Packaging

Hugh E. Lockhart
Major professor

Date November 30, 1998

The Effect of Subject Age on Legibility

By

Laura Bix

A THESIS

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

MASTER OF SCIENCE

School of Packaging

1998

Abstract

The Effect of Subject Age on Legibility

By

Laura Bix

There is an instrument that objectively quantifies how easy, or difficult label elements are to read. The instrument, the polariscope, does this by measuring the amount of light a subject requires to read a particular element. The difficulty in using this instrument as a performance standard for legibility is that as people age, physiological changes occur in their eyes. Because of this reality, older subjects with the same visual acuity as a younger subject will require more light to see clearly.

Four age groups, 21-35, 36-50, 51-65, and 66-80 were tested using cards printed in two fonts, Times-New Roman and Avant Garde. Each age group contained 20 subjects who read three messages printed in each of the two fonts. Subjects saw each font/message combination three times, for a total of 18 responses per subject. All subjects had measured visual acuities of 20/20 or 20/30.

The effect of age and font were found to be statistically significant at α=.01. Message was found to be significant at α=.05. The effect of font and message were not practically significant. The effect of age was overwhelming when compared to that of message and font.

Because visual acuity was confounded with age, subjects with visual acutities of 20/30 were excluded when a second ANOVA was conducted. The effect of age was still found to be significant at α=.01 and was much greater in magnitude than the effect of visual acuity.

Dedication

This thesis is dedicated to the communities of New Market, Iowa; Westerville, Ohio and East Lansing, Michigan who gave their time freely for nothing more than the chance to improve the labeling of medication bottles.

Acknowledgements

I would like to take this opportunity to officially thank my major professor, Dr. Hugh Lockhart. Dr. Lockhart was tirelessly available and supportive through this and many other endeavors. He is a patient and helpful person who always listens to student input and he never runs out of red ink.

I would also like to thank my committee members, Dr. Susan Selke and Dr. Dennis Gilliland. Like Dr. Lockhart, Dr. Selke is continuously available and always eager to help students in any way that she can. Dr. Gilliland has to be one of the most patient people that I have ever met. His steadfast efforts helped me to overcome my tremendous fear of statistics (somewhat), and even helped me to appreciate its power (somewhat).

To Dr. Diana Twede for the financial support and endless opportunities she has provided me during my education at the School of Packaging. She has not only taught me about Packaging, but also about the world through sharing her various experiences and thoughts.

"A teacher affects eternity; (s)he can never tell where his/her influence stops."
-Henry Brooks Adams

I cannot mention statistics without thanking my friend, Ben-rong Chen who came running anytime the going got rough. I am glad that I met this wonderful person both for his statistical prowess and his friendship.

Thank you to my good friend Sang-Jin Lee for his help with my late night computer snafus, and his ability to fix them. This thesis would have been completed on a typewriter if not for him.

To my many other friends at the School of Packaging who provided me shoulders to cry on, kept me company during many different trips to the library, provided rides when my car was out of commission and helped with a variety of home repairs. You are some of the most giving people I know. You have taught me to not take myself too seriously. Thank you.

I would like to thank my family, the Cades , the Bixes and the Simons, who came to my aid and helped me recruit subjects across the US. I would not have accomplished anything in my life had it not been for your undying support both emotionally and financially. (I'll try to stop asking soon).

Finally, to my husband who was there when the computer crashed and the statistics did not go well, when things needed to be proof read or when the books needed to be picked up or returned. Thank you for cleaning when people were coming over, or making dinner when I could not. You are my best friend.

"No matter what accomplishment you make, somebody helps you."
-Althea Gibson Darben

Table of Contents

List of Tables

List of Figures

Introduction

Many people find the labels of over-the-counter (OTC) medications difficult to read. This is a serious problem because over-the-counter drug labels carry important information such as warnings, directions, indications and ingredients. The seriousness of this issue is amplified for several reasons including: the aging of the population, the trend to move ethical (prescription) medications to OTC status and the popularity of self-diagnosing and self-medication as a health care option.

With age comes a propensity for certain diseases; as a result, many people need medication late in their lives. The U.S. Food and Drug Administration (FDA) (1) estimates, "the elderly comprise 12 to 17 percent of the population but consume about 30 percent of all medications." As the population ages, this percentage will increase; FDA projects that the elderly will consume 50 percent of all medications by the year 2000.

Another age-related problem is ocular degeneration. Degeneration of the eyes can make it difficult for older adults to effectively use information provided on OTC labels. A survey conducted by Weymouth (2) used subjects age 40 and older to gain an understanding of the relationship between visual acuity and age. Weymouth found that 92.6% of the 44-49 year old group maintained visual acuity of 20/20 or better. The percentage of people with 20/20 vision decreased in each age group; only 6.1% of the oldest group tested, 80 years and older, maintained a visual acuity of 20/20 or better.

The issue of legibility in OTC drug labels is further complicated by the fact that, "in recent years, more potent drugs have been switched from prescription to OTC status" (1). Examples provided by the FDA include cimetidie, naproxen sodium, ketroprofen, nicotine polarcrilex, nicotine transdermal system and minoxidil topical. When classified as OTC medications, safe and effective use depends on the appropriate application of written information; the consumer must be able to effectively read and understand OTC drug labels. Before a prescription drug can be reclassified as an OTC drug it must meet several criteria which include writing label instructions so that consumers can understand and follow them safely and effectively (3). Despite this requirement, even FDA (1) admits that not all OTC labels are easily read by the public, stating, "consumers often have difficulty finding, reading, and understanding this (OTC) labeling information."

The move of prescription drugs to OTC status combined with the rising cost of healthcare makes self-medicating an attractive alternative. The FDA (1) estimates that currently consumers are four times more likely to take an OTC drug for a health problem than they are to see a physician.

The FDA has not been the only government agency to address the importance of the ability to read and interpret labels. The Department of Transportation, the Consumer Products Safety Commission and the Environmental Protection Agency all regulate label legibility relating to labeling of hazardous materials (3)(4). FDA (5) currently regulates the labeling on ampoules, which are containers typically used for parenteral drug products. They have also proposed a rule (1) that would regulate the legibility of OTC medication labels if it is implemented.

The proposed rule was published in the Federal Register on February 27, 1997. It attempts to improve OTC drug labels by requiring: a specific font and font size, standardized headings and subheadings, bullet points, pictograms, specified words and an exact format for the required information. The proposed rule is inadequate because it does not measure the legibility of the message; it simply requires a specific font and font size on labels.

There are many critics of FDA's latest proposal. A study conducted at the New England College of Optometry (2) reports, "Type size alone may not be responsible for poor readability. Other factors that may be contributing to this difficulty include letter and line spacing, letter contrast, print and background color, and type style." The study concluded "horizontal letter compression had a greater effect on readability than vertical letter height."

An experiment conducted at the Michigan State University School of Packaging in July 1997 (6) also suggests that more factors influence legibility than font size and type. A message made from 4.5 point type with a black on white contrast was more easily read than the same message printed in 6 point type with a yellow on red contrast. If the proposed rule were implemented the 4.5 point type would be in violation, while the 6 point type would comply. The aforementioned results suggest that specifying a font with a minimum size requirement does not guarantee that messages will be legible. These results were obtained using an instrument known as the polariscope.

The polariscope is an instrument that objectively measures how easy, or difficult, label elements are to read. It does this by varying the amount of light required for a subject to read different label elements. One difficulty in using this instrument as a performance standard for legibility is that as people age, the physiology of their eyes changes so more light is required to see with the same clarity as their younger counterparts.

The hypothesis of this study is that the age of a subject has an effect on legibility readings measured by the polariscope. It may seem intuitive that with age more light is required, but an experiment designed to test this hypothesis using the polariscope is important if it is to be used as a performance standard for OTC drug labels.

Literature Review

Previous Work

Data presented in this thesis builds on the theses published by Mutune Wa Gitau(4) and James Pietrowski(7). Both did work under the direction of Dr. Hugh Lockhart using the polariscope.

Gitau (4) tested twenty subjects, ranging in age from 18-54. The legibility of instructions printed on 8 child resistant caps was measured using the polariscope and DMR (Design and Market Research Visibility Meter). An expert panel comprised of 10 professional printers was asked to evaluate the legibility of the same eight caps using a rating system of 1-10, the standard way to evaluate legibility in industry. High correlation was found when subject response was compared with expert response. A correlation coefficient of 0. 950 resulted when subjects with a visual acuity of 20/20 were tested and 0. 951 was the result when subjects with a visual acuity of 20/30 were tested. Bivariate plots comparing the polariscope and the DMR were also highly correlated. A correlation coefficient of 0.911 corresponded to subjects with visual acuities of 20/20 and a comparison of the two machines using subjects with 20/30 vision resulted in a correlation coefficient of 0.905.

Work done by Gitau (4) showed high correlation between accepted measures of legibility and the two instruments. In addition, his statistical analysis showed all main effects, message, visual acuity, contrast and type size, to be significant at α=. 05. Statistical analysis did not include the effect of age. However, he did make an observation regarding age, indicating that a "regular decrease in legibility with age in each visual acuity level suggests the possibility of an age effect on the legibility measurements."

Like Gitau, Pietrowski (7) correlated experts' ratings to subjects' results measured by the polariscope. Nine messages in two fonts were printed in three sizes, for a total of 54 cards. Correlation between polariscope readings and the expert panel was 0.984. Pietrowski found significant effects of visual acuity, age, type size and font.

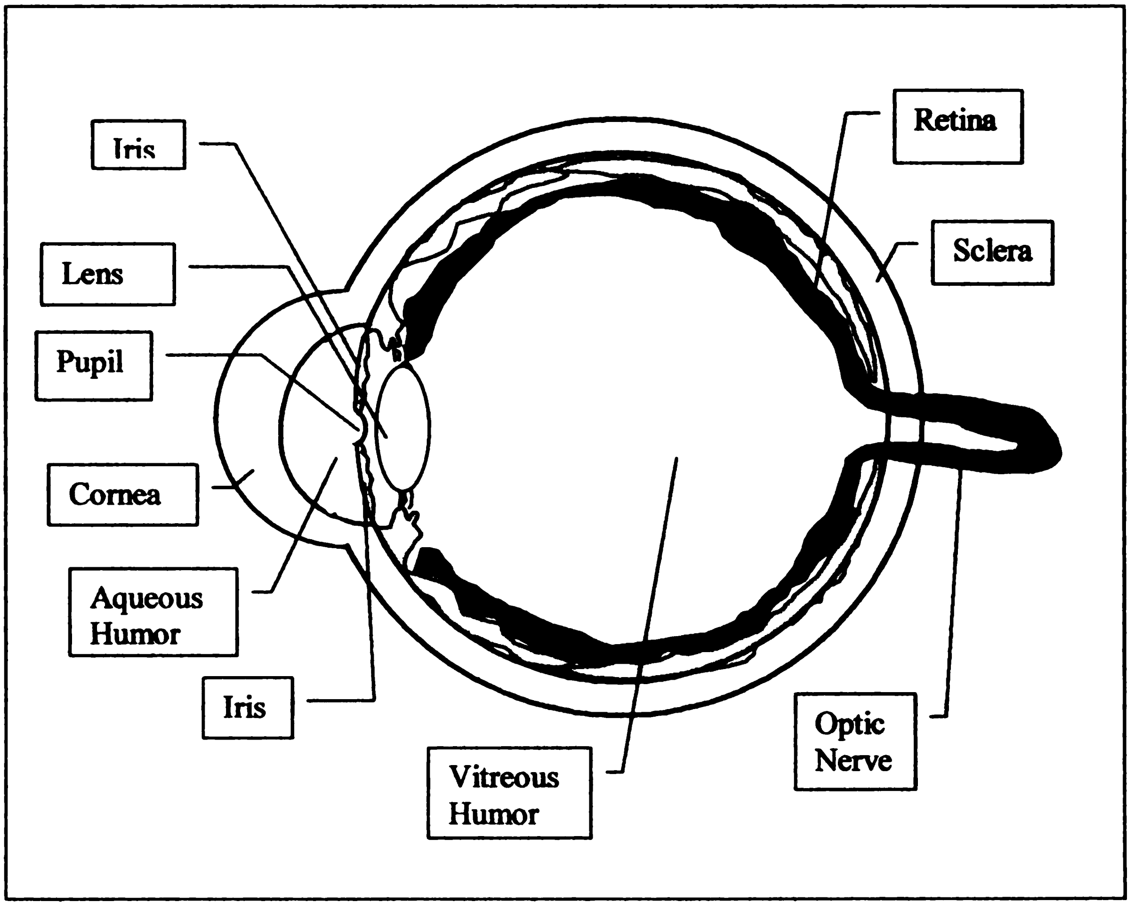
Both Gitau (4) and Pietrowski (7) suggested that age had a significant effect on the measurements obtained with the polariscope. Gitau did not test statistically for an effect of age, but anecdotally suggested its presence and recommended that it be further investigated. As a result of Gitau's observations, Pietrowski statistically analyzed the effect of age on legibility, finding it significant. Although Pietrowski tested for an effect of age, his experiment was not designed specifically to examine its effects. Specifically, he did not recognize that the effect of visual acuity is likely to be confounded with the effect of age.

This study builds on Pietrowski's (7) findings and Gitau's (4) observation that age has a significant effect on the degrees of rotation required by a subject to read a given message. It examines the separate effects of age and visual acuity, and suggests that experiments must be carefully designed so that these two effects are not confounded. This is something that neither of the previous two theses suggested. The idea that the aging process creates the need for increased light, even though the ability to focus remains sharp, probably seems intuitive to many readers. However, the effects of aging and visual acuity must be examined before the polariscope can be used to establish a performance standard for the legibility of OTC drug labels.

The Eye

According to Maryann Kelly (8) the eye undergoes several structural changes during the aging process. Kelly indicates that changes in the pupil, cornea, lens and retina (see Figure 1) combined with an increased propensity for ocular disease contribute to diminished visual function and the requirement for additional light in older adults.

FIGURE 1
THE STRUCTURE OF THE EYE (9)



The sclera and the cornea form the outermost portion of the eye. The sclera is a fibrous coating that protects and shapes the eye, providing an anchor for muscles. The cornea bulges from the sclera and is transparent, allowing light to enter the eye (9). As light enters the eye, it is refracted by the cornea, which provides approximately 75% of the eye's focusing power (8); the light then passes through the aqueous humor to the lens, where it is refocused and passed through the vitreous humor, a clear fluid, to the retina (9).

According to Leopold, as quoted by Kelly, the cornea's most significant change due to aging is a "flattening", which occurs after age 65 resulting in astigmatism. Degenerative changes also occur in the sclera, which result in a loss of clarity that affects the quality of vision of older adults.

Aging also affects the retina and the lens of the eye. As aging takes place, the retina begins to thicken and yellow. Retinal changes, combined with a loss of elasticity, thickening, and yellowing of the eye's lens, reduce the short-wave lengths of light (the blue color) entering the eye. The resultant effect is that colors such as violet, blue and green are filtered out, reducing the contrast sensitivity of the eye and causing glare which interferes with the visual image. In addition, a gradual loss of lens transparency can affect clarity of vision in older adults.

Visual quality of older adults is also affected by changes in the pupil. "The iris is unable to dilate as much as in youth under all light conditions, but this is especially evident and troublesome under dark-adapted conditions" (8). As a result of the iris's inability to dilate, the size of the pupil decreases with age. Kelly quotes Marmor and Sullivan as indicating that the average pupil size at night in a 20-year old is 3.3cm, compared with 0.2 cm in an 80-year old. This creates a kind of chain reaction. The iris is unable to dilate, keeping the pupil small and, ultimately, decreasing the amount of light reaching the retina, which, as a result of age-related thickening, is not as sensitive as in younger subjects.

Not only do older adults have to contend with physiological changes that result in diminished visual function, they are also more likely to develop certain ocular pathologies. Cataracts, glaucoma, senile macular degeneration (SMD), diabetic retinopathy, and presbyopia, a decrease in the ability to focus, are diseases that are more prevalent in the elderly (8)(3).

The effects of these conditions may be more serious in the older population because as people age, the frequency of their visual exams decreases. Data from the National Center for Health Statistics indicate, "Vision care examinations for the retired population decrease by 50% from pre-retirement levels, probably due to financial concerns" (8).

The many physiological changes of the eye, in combination with a number of other factors contribute to the potential for mismanagement of OTC drugs in the elderly population.

Self Medication and Drug Mismanagement

Holt (10) estimates that there are as many as 600,000 non-prescription drug products available to American consumers; these products are manufactured by 12,000 firms and involve 700-800 active ingredients. He suggests that the large number of products available combined with the cherished American ideals of independence and freedom lead many consumers to treat themselves with OTC medications. These ideals combined with the increase in OTC drugs available and the cost savings they provide have increased their popularity. FDA (1) estimates that consumers are four times more likely to medicate themselves than see a physician for a health problem.

The frequent use of nonprescription drugs is complicated by the fact that many people supplement prescribed medications with self-prescribed products (10). This is of particular concern in aging adults, who are more likely to have multiple conditions requiring medication, frequently turn to over-the-counter medications due to financial concerns, and have to contend with significant changes in the eye's physiology and any resultant difficulties with reading information on the labels.

A study published in the Journal of The American Optometric Association (2) indicates that the elderly do have difficulty reading OTC labels. The study tested the legibility of three OTC labels using 92 subjects age 60 and older with visual acuity of 20/60 or better. Using the "rule of 1000", researchers were able to estimate the visual acuity required to read certain labels under controlled lighting conditions. The "rule of 1000" estimates the acuity required to read a given label. According to the rule, the number of letters and spaces in one inch are counted. This number is then divided into 1000. If the result is larger than the Snellen denominator of a subjects' near visual acuity, then he or she should not need extra magnification. (The Snellen denominator is simply the denominator of their visual acuity. If a subject had a visual acuity of 20/40, their Snellen denominator would be 40). Conversely, if the number is smaller than the Snellen denominator, the subject may need extra magnification.

The "rule of 1000" estimates indicated that several of the labels tested would require subjects to have a visual acuity of 20/20 or better. Although 20/20 vision is frequently known as "normal" it is more the ideal than reality for the population over 60 years of age (3). The study (2) found that 76.1 percent of subjects had visual acuities worse than 20/20. "If three-fourths of the 60 and over population cannot read 20/20, yet are given labeling with print that is 20/20… it is expected that many of them will not be able to read such a label." Researchers (2) did find that "a significant portion of the elderly population cannot adequately read the print on certain OTC medication labels due in part to small vertical type size and high degrees of horizontal letter compression."

A study of drug-implicated hospitalization (11) provides insight into the potential effects of the inability to effectively use drug labels, and demonstrates that this is a concern in the elderly population. For three years a research team from the University of Florida investigated all admissions to their Teaching Hospital that were the result of drug-induced illness, excluding suicide and drug abuse. Patients were excluded from the study if minor drug reactions were present, but the reaction was not the major reason for admission. 177 out of 6, 063 admissions, or 2.9%, were due to drug induced illness. "Proportionately fewer admissions for drug-induced illness were observed in patients under 60 years of age, but between 61 and 80 years of age, proportionately more patients were admitted because of drug reactions. Only in the 71 to 80 year old group was there a significantly ($p \leq .01$) greater number of admissions due to adverse drug effects".(11) In 18 percent of the 177 cases, an over-the-counter drug was implicated, and more than 6% of patients in the study died.

"Elderly patients account for 39% of all hospitalizations and 51% of deaths from drug reactions…Mismedication is thought to be a significant problem with this population" (10). Researchers in the study presented in this document bore witness to the fact that drug mismanagement among the elderly population is a reality as many test subjects indicated that they could not read the directions or warnings on many of their medications. The gravity of the problem was obvious when one 80 year old test subject indicated, "I can never read those damn things (OTC labels). So I just take two." It is time that a performance standard for legibility is applied to the vital information contained on the labels of OTC medications.

Materials and Apparatus

Legibility Cards

Eight cards, created using Microsoft Word, were used in this study. The cards were labeled one through eight so that they could be easily identified. Cards one through four were printed with 14 point Times New Roman, a serifed font, and cards five through eight were printed with 14 point Avant Garde, a sans-serif font. (See Table 1) In this study, the Times New Roman is referred to as "Font 1" and the Avant Garde is referred to as "Font 2".

Four messages were used in this study. Each message was assigned a number. Cards one through four contained the message numbers one to four, respectively, in font one. Cards five through eight contained messages one through four, respectively in font 2. (See Table 1) The messages had been used in previous work at the School of Packaging. Previous ANOVAs conducted on the messages found them to not be statistically different at level $\alpha=0.05$.

After messages were printed using a Hewlett Packard Ink Jet printer, they were cut into 3 inch by 4 inch cards, with messages centered horizontally on each. All messages were printed using a black on white contrast. Black on white was chosen because it provides a high degree of contrast and is generally recognized as easy to read.

TABLE 1

Card #	Font #	Message #	Font	Message- As it appeared on the card
1	1	1	Times New Roman	It may help most of them to work today. She works in this club after midnight. The order to go will be done after two.
2	1	2	Times New Roman	She works in this club after midnight. The order to go will be done after two. There will be some sugar in the kitchen.
3	1	3	Times New Roman	The order to go will be done after two. There will be some sugar in the kitchen. Here is a copy of lunch hours for today.
4	1	4	Times New Roman	There will be some sugar in the kitchen. Here is a copy of lunch hours for today. From here to there flowers cannot grow.
5	2	1	Avant Garde	It may help most of them to work today. She works in this club after midnight. The order to go will be done after two.
6	2	2	Avant Garde	She works in this club after midnight. The order to go will be done after two. There will be some sugar in the kitchen.
7	2	3	Avant Garde	The order to go will be done after two. There will be some sugar in the kitchen. Here is a copy of lunch hours for today.
8	2	4	Avant Garde	There will be some sugar in the kitchen. Here is a copy of lunch hours for today. From here to there flowers cannot grow.

Dow Corning Ophthalmics Card

Prior to testing, the visual acuity of each subject was measured using the Dow Corning Ophthalmics near point visual acuity card. (See Figure 2) The instructions on the card state that subjects are to hold the card "16 inches from their eyes in 'good light'" (12). Subjects were instructed to wear their prescribed lenses, such as bifocals or reading glasses, while their near vision was tested.

Visual acuity measurements from the Dow Corning Ophthalmics near point visual acuity card utilize a standard format for acuity results that was developed by Snellen. The Snellen notation is based on a 20-foot distance. The denominator indicates the distance of the smallest numbers seen by the test subject. The numerator is the distance at which an average person sees the same numbers. In other words, scores of 20/20, 20/40, 20/200 etc. indicate that the subject being tested at 20 feet can just discriminate letters that a person with normal vision, the average person, can see at 20 feet, 40 feet, 200 feet and so forth (8).

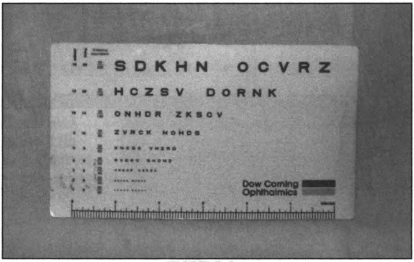


FIGURE 2
(card is reduced from actual size)
DOW CORNING OPHTHALMICS NEAR POINT VISUAL ACUITY CARD

Pseudo-Isochromatic Plates for Testing Color Perception

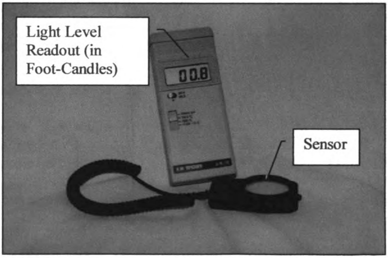
Subjects were tested for color perception using pseudo-isochromatic plates manufactured by Richmond Products. Subjects examined 15 color plates; responses to the plates were recorded and tabulated. The results of the color perception test were not used in the statistical analysis.

A.W. Sperry Light Meter (Model SLM-110)

A light meter manufactured by A.W. Sperry was used to ensure consistent light levels inside the polariscope. (See Figure 3) To measure the light level inside the polariscope, the sensor was placed on the lower easel and the lid was closed. The light inside the polariscope was adjusted using a rheostat until the light meter reached a level of 25 foot-candles ± one foot-candle.

To measure the level of light in the room, the sensor was placed on the table, facing the ceiling, to the left of the polariscope. The light level at each test site was recorded but could not be adjusted. All testing was conducted under florescent room lights in an attempt to maintain consistency.

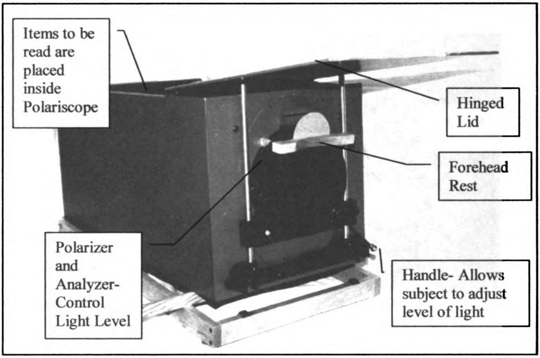
FIGURE 3
A.W. SPERRY LIGHT METER AND SENSOR



The Polariscope

The polariscope has evolved since its creation in the 60s. Initially, the polariscope was developed as a way to measure label impact; it was primarily a marketing device. Dr. Hugh Lockhart at Michigan State University's School of Packaging recognized that the instrument was not being used to its potential and began developing it as a way to measure legibility. Pietrowski (7) constructed the instrument used in this study, under the direction of Dr. Lockhart, in 1993. (See Figure 4)

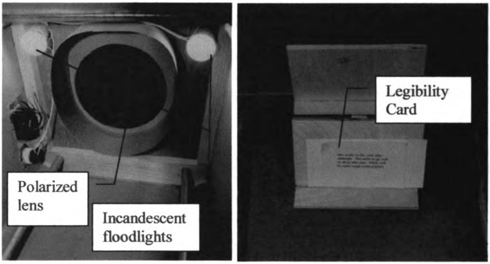
FIGURE 4
POLARISCOPE- OUTSIDE



The polariscope uses polarizing filters to control the amount of light that reaches a test subject's eyes. Test subjects control the amount of light reaching their eyes by rotating the handle on the right side of the machine. The filters are Polaroid HN22 Linear Polarizing Filters that are .030 inches thick. Pietrowski (7) used the HN22 filters because they had a uniform level of light transmission throughout the portion of the spectrum to which the eye responds, 440-750 nm wavelength.

The item to be tested, in this case a printed card, is placed inside the polariscope on one of two easels. (See Figure 5) The lower easel was used for this experiment. Two 25-Watt Sylvania floodlights illuminate the test material. (See Figure 5) Power to the floodlights, and, ultimately, the light level, is controlled by a rheostat manufactured by Powerstat.

FIGURE 5
POLARISCOPE- INSIDE



Two incandescent 25 watt Sylvania Floodlights (left) illuminate the material to be tested, which rests on one of two easels inside the polariscope (right).

The card is placed on the lower easel approximately 17.5 inches from the subjects' eyes. This distance was chosen because it is within the normal range for reading distance (3). As the subject rotates the outer filter, more and more light is allowed through the pair of filters. Each subject begins with the legibility index at zero, the point where no light passes through the filters, and can continue rotating the lens until maximum light is allowed through, a legibility index of ninety degrees. The more difficult a message is for the subject to read, the more light it requires. Thus, messages that are difficult to read result in higher responses, recorded in degrees of rotation, also called the legibility index, than messages that are easily read.

Subjects' Demographics

Initially, four age groups, 21-35 (age group 1); 36-50 (age group2); 51-65 (age group 3) and 66-80 (age group 4), were targeted for this study. Because of the availability of subjects, the subjects in each age group were inhabitants of different locations. Age group 1 was primarily composed of students, graduate and undergraduate, that attended Michigan State University. Age groups 2 and 3 were residents of Columbus, Ohio and age group 4 consisted of people living in the community of Clarinda, Iowa. (In the analysis portion of this paper the data obtained from the Iowa group is reported as "age group 5").

Because the effect of geography was confounded with age for the aforementioned design, it was decided that a fifth group would be added and examined to determine any effect of geography. The fifth group, age 66-80, were from mid-Michigan. (In the analysis portion of this paper the data obtained from the 66-80 year old Michigan group was reported as "age group 4"). The responses provided by this age group were compared against responses provided by age group 5 to determine whether a significant effect of geography existed. The results of this comparison are covered in the "Results" section of this paper and are based on the analysis contained in "Appendix F".

Statistical Methods

The responses were analyzed using a mixed model Analysis of Variance (ANOVA). The factors "age group" (5 levels), "font" (2 levels) and "message" (3 levels) are crossed and the factor "subjects" (20 per age group) is nested within age group. Message and subject are treated as random effects in the reported analyses (Appendices D, E and F). Each subject was presented with each font-message combination three times.

Examination of the residuals shows that the normality assumption is supported when the raw responses are replaced by their logarithms. The p-values for tests of significance are based on F and approximate F-statistics derived from ANOVAs of logged data (Appendices D, E and F). Satterthwaite's method (13) for determining approximate F-statistics and their degrees of freedom is used. The statistical analyses were performed using Minitab Version 11.12.

The effect of "age group" is of most interest in this research. Of course, age group 4 (66-80, Michigan) and age group 5 (66-80, Iowa) are distinguished by location rather than age. The other three levels are age group 1 (21-35, Michigan), age group 2 (36-50, Ohio) and age group 3 (51-65, Ohio). The Appendix D analysis uses all five groups. The Appendix E analysis uses the responses from subjects with 20/20 visual acuity only. The appendix F analysis uses the responses from age group 4 and age group 5 only.

Procedures for Data Collection

In this experiment, five groups were tested using eight cards that contained two fonts and four messages. Subjects were recruited through friends, family and local community groups. Testing was conducted in four different locations. All tests were conducted in rooms lit by florescent lights in an attempt to maintain consistency between locations.

Subject Orientation

Prior to testing, subjects were provided with a brief one-on-one orientation with the researcher. The orientation began with the researcher asking subjects if they had ever had any difficulty reading the instructions provided on OTC drug labels. Comments or anecdotes that were relevant to the study were recorded without attribution to the individual.

Following the discussion of difficulties with OTC labels, the researcher explained,

> "The polariscope is an instrument that quantifies how easy, or difficult a label is to read. It does this by measuring the amount of light a subject requires to read a given message. The harder a message is to read the more light is required; the easier it is to read, the less light is required. If you choose to participate in this study you will be asked to fill out some information regarding your education, eyewear and age. This information will be anonymous; your name will not be recorded on any documents. Your visual acuity and color perception will be measured. After we have collected all of that information, you will be asked to read a total of 20 cards that are in the polariscope. If you normally use corrective eyewear to read you should use it for this experiment. You will read the cards by looking into the round screen on the front of the polariscope. Notice the headrest above the round screen; many subjects feel more comfortable if they lean against the headrest during the test. You may choose to use the headrest or you may sit back from it; do whatever makes reading the card the most comfortable for you. As you look through the round screen rotate the handle on the right side of the machine until the first point where you are able to easily read all the words on the card without straining your

> eyes. Testing will not take longer than 20 minutes. It is important to remember that this is a test of the printed cards and not of your eyes. There is no need to worry about how your results compare with the results of other test subjects. We are concerned with how the cards compare with one another."

In order to protect the test subjects' rights, subjects were asked to review and sign a written consent form. Subjects could either sign that they accepted or declined the request to participate in this study. (See Appendix A)

Collection of Subject-Related Information

After the subject had signed the consent form, a data-recording sheet was used to record data regarding subjects and the results of their legibility test. (See Appendix B) These sheets identified the order of the subjects' appearances, chronologically recorded as "subject number". The first subject measured in each age group was labeled subject #1; the second in each group was labeled two, etc.

After background information (eyewear, age and education) was recorded, subjects were given the Dow Corning Ophthalmics Near Point Visual Acuity card. (See Figure 2) While seated in front of the polariscope, they were asked to hold the card approximately 16 inches in front of them and read the smallest print that they could read. The researcher was prepared to suggest adjustment in the event that subjects positioned the card substantially off the 16 inch requirement. This was unnecessary, however, because all subjects who had difficulty gauging the distance requested help. Each subject's Snellen visual acuity was recorded on the data-recording sheet. Data was collected from all subjects willing to participate in this study; however, only subjects with measured visual acuities of 20/30 or better were used for analysis purposes.

After each subject's visual acuity was measured, their color perception was tested using a book of pseudo-isochromatic plates. Subjects were instructed, "Look at each page in this book. If you see a number on the page, tell me what number it is; if you do not see a number just tell me that you do not see one. Do not be alarmed if you do not see a number on each page." The subjects' responses for each of the fifteen plates were recorded on their data-recording sheet. These responses were later tabulated according to the directions at the back of the test booklet but were not used in the statistical analysis portion of this paper.

Data Collection

The legibility cards were tested using the polariscope after the subject's background information was recorded completely. It was explained again that this was a test of the legibility cards, not their eyesight. Subjects were reminded to rotate the handle to the right of the machine to the first point that they could easily read the message on the card inside of the polariscope. The room lights remained on during testing.

In a previous study (14), conducted in December 1997, an analysis of residuals revealed that primary readings tended to be higher than readings that followed. This suggested that subjects go through an "adjustment period" as they get used to using the polariscope. As a result of this information, it was decided that a set of "dummy cards" would be used. The "dummy cards" allowed subjects to become used to the polariscope without affecting the data collected. Subjects always read the "dummy set" prior to reading any other cards.

The "dummy set" consisted of two cards. Card # 4 was message #4 printed in font #1, Times New Roman. Card #8 was message #4 printed in font #2, Avant Garde.

(See Table 1) The order which these two cards appeared rotated, i.e. if subject one saw card #4 first and card #8 second, then subject two saw card #8 first and card #4 second. Because these cards were intended to provide subjects with an adjustment period, legibility indices from subjects when viewing cards #8 and #4 were recorded, but not included in the statistical analysis.

The remaining 6 cards, numbers 1, 2, 3, and 5, 6 and 7 were tested and legibility indices from these cards were used in the analysis section of this paper. Each card carried a message/font combination that differed from the others. (See Table 1) Each subject saw each card a total of three times, for a total of 18 responses per subject. Subjects were not told that cards were repeating. The legibility index from each card, including those generated by the "dummy set", was recorded on the data-recording sheet.

When time permitted, readings of the light level inside and outside the polariscope were taken between subjects. These readings were recorded on the previous subject's data-recording sheet. If necessary, adjustments were made to maintain a light level of 25 foot-candles ± one foot-candle inside the polariscope.

Results

Table 2 summarizes the average legibility index for the five age groups for all treatment combinations, providing useful information for interpreting the ANOVA results. (See Appendix C for Descriptive Statistics: See Appendices D, E and F for ANOVA Results)

TABLE 2

Average Legibility Index for Each Age Group and Treatment Combination						
Font*	Message	Age Group 1	Age Group 2	Age Group 3	Age Group 4	Age Group 5
TNR	1	16.4	21.1	25.2	34.5	48.2
TNR	2	16.4	21.6	25.0	33.9	45.8
TNR	3	16.4	20.6	25.2	34.5	45.8
AG	1	15.6	19.9	23.2	32.2	44.4
AG	2	15.2	19.9	22.9	31.9	45.6
AG	3	15.8	19.3	23.2	31.1	44.5
Average Legibility Index		16.00	20.4	24.1	33.0	45.7
*TNR= Times New Roman AG = Avant Garde						

Across all treatment combinations the legibility index, measured in degrees of rotation, increases with increasing age. In order to account statistically for the effects of age, font, message, subject and acuity, the data was analyzed using several techniques. (See Appendices D, E and F) To ease interpretation of the results, graphs presented in the "Results" section were created using the raw legibility index. The statistical analyses presented in Appendices D, E and F use the natural log of the legibility index. Since the calculations of significance in the ANOVA assume a normal distribution, the natural log of the legibility index was used to normalize the data.

The ANOVA (Appendix D) conducted using data from all 5 age groups revealed the effect of both age and font to be highly significant at α=0.01. The estimated variance component for the main effect of subjects nested within age group is .143 (Appendix D). Not surprisingly, the subject to subject variation is statistically significant. No other effects were found to be significant at level α=0.01. (See Table 3)

Although message is not significant at α=0.01 it is statistically significant at α=0.05. Despite this fact, it is not practically significant. The largest difference in the grand average legibility index for the three messages is 0.4 degrees. (The grand average is calculated by averaging all legibility indices collected for a single effect. For example, the grand average for message number one is calculated by averaging responses from all subjects for cards 1 and 5). By contrast, the largest difference in the grand averages for each age group is 29.7 degrees. Clearly, the effect of age is overwhelming when compared to message and font. (See Tables 3 and 4)

TABLE 3

ANOVA Results- Using the Restricted form of the Mixed Model and the Natural Log of the Legibility Index (Appendix D)	
Effect	Calculated p-value
Age	**$4.7 * 10^{-13}$**
Font	**$2.9 * 10^{-4}$**
Message	0.035
Bolded effects are significant at α=.01	

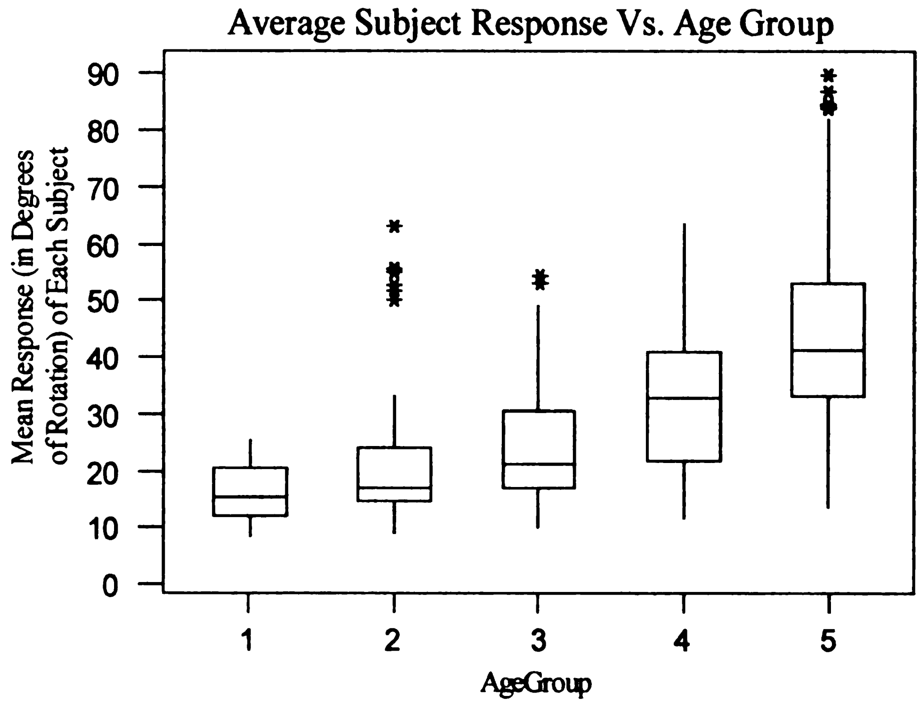
These results are qualitatively the same as those found in the analysis of other models. When only age groups one through four are used, qualitatively the results are the same. The results are also qualitatively the same as those listed in Table 3 when message is treated as a fixed effect.

TABLE 4

	Treatment	Average Legibility Index	Largest Difference in Grand Averages
Font	Times New Roman	28.7	1.7
	Avant Garde	27.0	
Message	Message 1	28.0	0.4
	Message 2	27.8	
	Message 3	27.6	
Age	Age Group 1	16.0	29.7
	Age Group 2	20.4	
	Age Group 3	24.1	
	Age Group 4	33.0	
	Age Group 5	45.7	

Box and whisker plots created with the grand average across all treatments for each subject graphically show the magnitude of the differences between age groups. (See Figures 6 and 7) The lowest legibility indices increase slightly with increasing age, as indicated by the low end of the "whisker". As age increases, the spread of the data increases and the highest values, as well as the medians (and means) increase substantially. This increase is so large that the difference in the average legibility index of age group five is 29.7 degrees higher than age group one's average. (See Table 4)

FIGURE 6



The bottom line (horizontal) of each plotted box represents the 1st quartile of the data in each age group; the top of the box represents the 3rd quartile of the same data set. The vertical lines off of the boxes, or "whiskers", are the lowest and highest legibility indices within the group that still remain within three standard deviations of the data set. The "outliers" are represented as asterisks.

FIGURE 7

Treatment Combinations Across all Age Groups

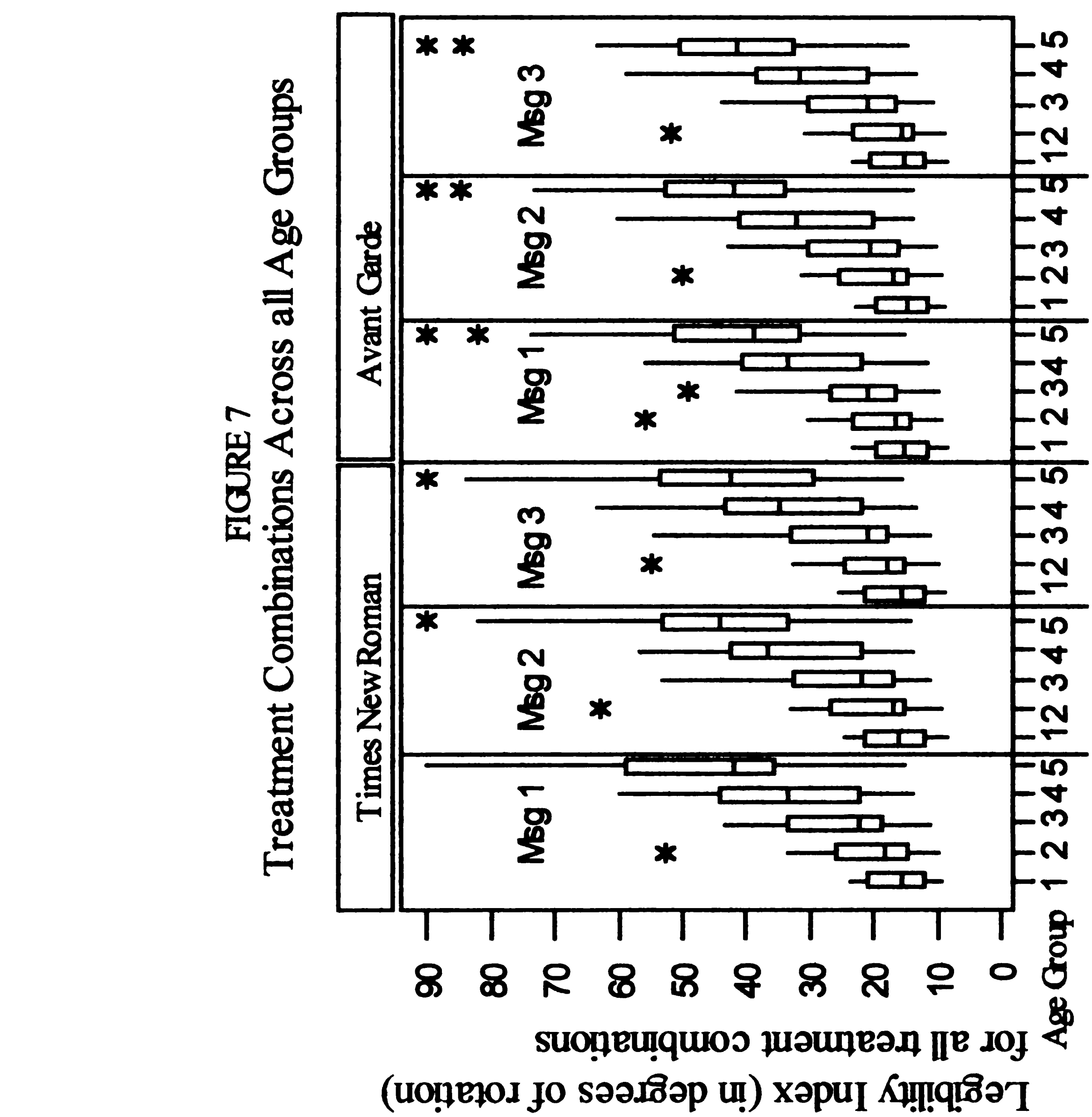


Figure 7 was created by averaging the three repetitions for each treatment combination and generating a box plot of the legibility indices for all subjects within an age group. Each card represents a combination of treatments. By grouping the data in this fashion, the effect of font, message or age can be examined. For instance, although font was found to be statistically significant at $\alpha=0.01$, by examining the graphs it becomes evident that the effect of font is minor when compared with the age effect.

Subjects with measured visual acuities of 20/20 and 20/30 were included in the analysis of data presented to this point in this paper. As mentioned previously, with age comes a decrease in visual acuity. This is evident in the data collected for this study. (See Tables 5 and 6 and Figure 8.)

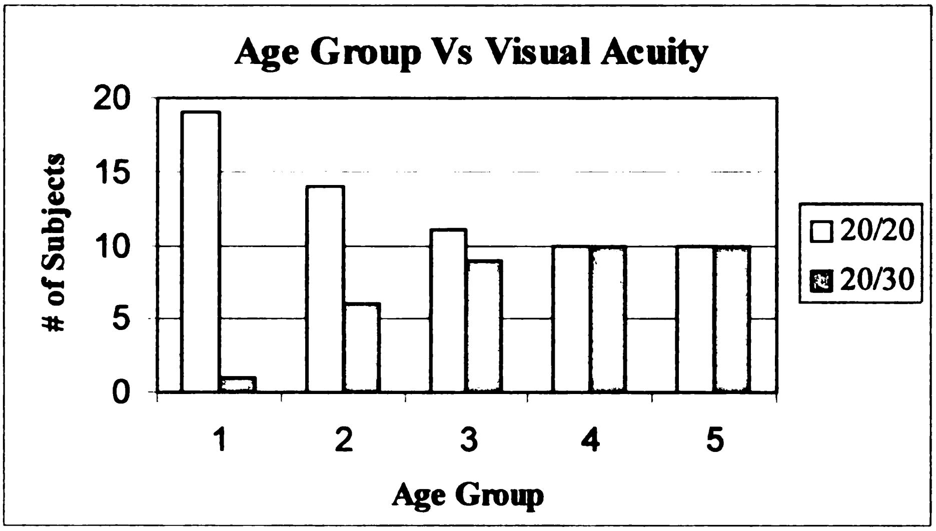


FIGURE 8

All graphs prior to Figure 8 used data from groups that included some subjects with visual acuities of 20/20 and other subjects with visual acuities of 20/30. Because of the trend of decreasing visual acuity with age, the effect of age and visual acuity were confounded in the data. In order to examine the effect of age independent of acuity, each age group was divided into two sub-groups, those with a visual acuity of 20/20 and those who were tested to be 20/30. The average legibility index for each sub-group was then calculated. Differences between groups and subgroups were examined. (See Tables 5, 6 & 7.) Box and whisker plots were created to examine differences visually. (See Figure 9.)

TABLE 5
Grand Mean of Legibility Index (Number of Subjects)

Measured Acuity	Age Group 1	Age Group 2	Age Group 3	Age Group 4	Age Group 5
20/20	16.2 (19)	19.4 (14)	22.8 (11)	29.9 (10)	43.5 (10)
20/30	11.7 (1)	22.8 (6)	25.7 (9)	36.1 (10)	48.0 (10)
Combined	**16.0 (20)**	**20.4 (20)**	**24.1 (20)**	**33.0 (20)**	**45.7 (20)**

An additional ANOVA was conducted using data collected from the first 10 people with measured visual acuity of 20/20 in each age group. All 20/30s were eliminated from this analysis. Despite the fact that no 20/30s were used in the analysis, the effect of age was still statistically significant at α=.01.(See Appendix E)

FIGURE 9
Box plots of Average Response Vs Age and Acuity

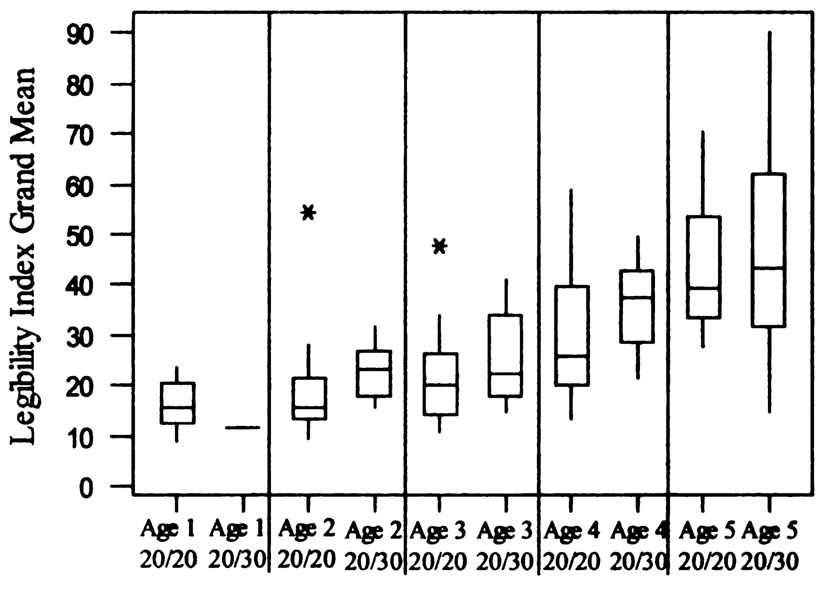


TABLE 6

Grand Mean of Legibility Index (Number of Subjects)

Acuity	Age 1	Age 2	Difference in Legibility Index Between Group 1 and 2	Age 2	Age 3	Difference in Legibility Index Between Group 2 and 3	Age 3	Age 4	Difference in Legibility Index Between Group 3 and 4	Age 4	Age 5	Difference in Legibility Index Between Group 4 and 5
20/20	16.2 (19)	19.4 (14)	**3.2**	19.4 (14)	22.8 (11)	**3.4**	22.8 (11)	29.9 (10)	**7.1**	29.9 (10)	43.5 (10)	**13.6**
20/30	11.7 (1)	22.8 (6)	**11.1***	22.8 (6)	25.7 (9)	**2.9**	25.7 (9)	36.1 (10)	**10.4**	36.1 (10)	48.0 (10)	**11.9**

*** There was only one individual with a visual acuity of 20/30 in age group 1.**

TABLE 7

Difference in Legibility Index (Average 20/30 Response-Average 20/20 Response within the Same Age Group)				
Age Group 1	Age Group 2	Age Group 3	Age Group 4	Age Group 5
-4.5*	3.4	2.9	6.2	4.5

*Average20/20 response was higher than the 20/30 response for age group 1. There was only one individual with a visual acuity of20/30 in this age group.

With the exception of age group one (which only had one subject with a visual acuity of 20/30) the average legibility index of subjects with measured visual acuities of 20/30 was greater than the average legibility index from subjects with measured visual acuities of 20/20 in the same age group. This was the expected result. A Snellen notation of 20/20 indicates that a subject being tested at 20 feet can discriminate letters that a person with normal vision, the average person, can see at 20 feet; a Snellen value of 20/30 indicates that a subject being tested at 20 feet can discriminate letters that a person with normal vision can see at 30 feet.(8) Therefore, it is not surprising that subjects with 20/30 vision had a higher average legibility index than subjects with 20/20 vision.

As a rough measure of statistical significance for the difference in means of these two groups, a two-sample t-test was applied to residuals, resulting in p-value 0.11. The Kruskal-Wallis test resulted in a p-value 0.06. This suggests that acuity cannot be ignored in experiments designed to compare messages and fonts where small differences in effects are to be estimated.

The effect of visual acuity does not, however, have the same magnitude as the effect of age. Figure 9 indicates that, within each age group, the legibility index of the subgroup containing subjects with visual acuities of 20/30 tends to be higher than that of the subgroup containing subjects with 20/20 vision. However, when these differences are compared to the differences in legibility index between age groups, it is clear that age has a greater effect. As a result of these findings, age should be the central concern when designing experiments that use the polariscope to measure legibility.

As mentioned previously, four age groups, 21-35 (age group 1); 36-50 (age group 2); 51-65 (age group 3) and 66-80 (age group 4), were targeted for this study. Due to the availability of subjects, the age groups were inhabitants of different locations. Age group 1 was primarily composed of students, graduate and undergraduate, that attended Michigan State University. Age groups 2 and 3 were residents of Westerville, Ohio and a fourth age group (age group 5 in the data presented in this paper) consisted of people living in the community of Clarinda, Iowa. Because the effect of geography was confounded with age for the aforementioned design, it was decided that a fifth group would be added to determine the probability that geography had an effect on the results.

The fifth group, age 66-80, were from mid-Michigan. (Legibility indices from this group were listed as "age group 4" since their average age was younger than that of the Iowa group). The legibility indices provided by this group were compared against legibility indices provided by age group 5, people 66-80 from Iowa, to determine whether a significant effect of geography existed. The average legibility index of subjects in age group 5, the Iowa group, was higher than those of age group 4. This was true across all treatment combinations and both visual acuities. (See Tables 8 and 9)

TABLE 8

Comparing the Average Legibility Indices of Age Group 4 (Michigan 66-80) and Age Group 5 (Iowa 66-80)			
Font	Message	Age Group 4 (MI)	Age Group 5 (IA)
Times New Roman	Message 1	34.48	48.17
Times New Roman	Message 2	33.87	45.80
Times New Roman	Message 3	34.48	45.80
Avant Garde	Message 1	32.23	44.37
Avant Garde	Message 2	31.87	45.60
Avant Garde	Message 3	31.08	44.48
Grand Average		**33.00**	**45.70**

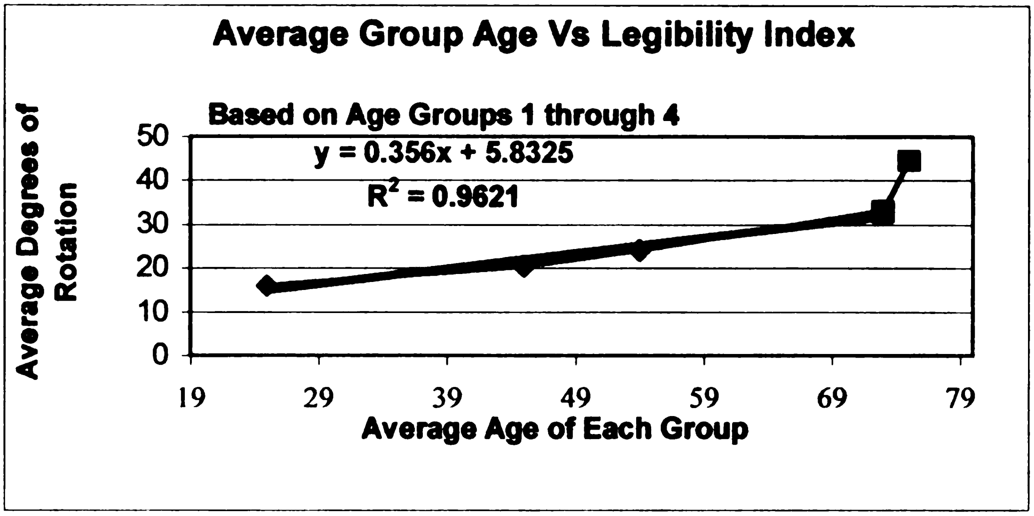
TABLE 9

Comparing the Grand Mean Response of Groups 4 (MI) and 5 (IA) (Indicates number of subjects)			
		Age Group 4 (MI)	Age Group 5 (IA)
Acuity	20/20	29.9 (10)	43.5 (10)
	20/30	36.1 (10)	48.0 (10)
	Combined	33.0 (20)	45.7 (20)

When the legibility indices of age groups 4 and 5 were compared using ANOVA it was determined that a statistically significant difference existed at α=.02. (See Appendix F) Because the age groups, as defined in this study, spanned 14 years, researchers investigated the specific ages of individuals contained in each group to determine if the aforementioned effect could be attributed to age, as opposed to geography. The average age of age group 4 was determined to be 73, while group 5 had an average age of 75. A nearly linear relationship was found in the first four age groups when a bivariate plot of average degrees of rotation versus the average age of subjects in each group was produced; a linear regression using only the first four age groups yielded an R^2 value of .9621. This relationship did not hold true when group 5 was added. (See Figure 10)

Readers should use caution when examining Figure 10. The figure is a bivariate plot of the average response across all subjects and treatments verses the average age of the entire group tested. The tendency of an average is to collapse the data. Readers are urged to recall that individual responses contained substantial variability.(See Figures 6 and 7)

FIGURE 10



Clearly, there are differences in the legibility indices of group 4 and group 5, despite the fact that there was only two years difference in the average age of the subjects contained in these groups. There are several possible explanations for this difference, which will be discussed in the "Conclusions and Recommendations" section of this report.

Conclusions and Recommendations

The hypothesis of this study is that the age of a subject has an effect on legibility readings measured by the Polariscope. All of the data collected for this study (See "Results" and Appendixes C, D, E and F) confirms this hypothesis. The effect of age proved to be more significant than any other effect, including visual acuity. At first glance this is a puzzling result, which can be better understood by re-examining the physiology of the eye.

"Visual acuity is the capacity to resolve the fine details of objects in the visual field under **ideal conditions** of brightness and contrast" (8). In other words, it is the ability of the eye to focus in bright light on items with sufficient contrast.

Declining visual acuity can be attributed to changes in the structures in the eye that are responsible for its ability to focus. According to Kelly (8), the cornea provides approximately 75% of the eye's focusing power. After light passes through the cornea, it travels through the aqueous humor to the lens, where it is refocused and passed through the vitreous humor to the retina (9). Changes that occur primarily in the cornea and lens during the aging process affect the eye's ability to focus effectively and result in decreased visual acuity. Some studies indicate that this decline begins as early as the second decade of life (8).

A key to understanding why an 80 year old with a visual acuity of 20/20 has a higher legibility index when using the polariscope than a younger subject with 20/20 vision lies in the phrase "under ideal conditions of brightness and contrast". As people age not only do they lose their ability to effectively focus, other physiological changes result in diminished light reaching the retina. Because tests for visual acuity are conducted under bright conditions, the effect of these changes does not register when a subject's visual acuity is recorded.

In this experiment the light falling on the message is low, only 25 foot-candles. The polarizing filters further reduce this level. Because the polariscope is dependent on varying the amount of light subjects use, the physiological changes that result in diminished light reaching the retina clearly manifest themselves in the data collected.

The amount of light reaching the retina is so important that Fozard et al (8) suggest that "the quality of retinal image is largely determined by the amount of illumination falling on the retina." This explains why two subjects with the same visual acuity in different age groups have drastically different results when using the polariscope. The effect of age is so dramatic that the difference in legibility index is greater than when two subjects of different visual acuities in the same age group are measured.

Age not only affects the legibility indices collected from individuals, it also affects the spread of the data collected. In general, as the age group increases, the difference between the maximum and minimum values in legibility index also increases. The box and whisker plot in Figure 6 shows this visually. (See Figure 6 and Table 10.)

TABLE 10

Difference in Maximum and Minimum Values for Each Age Group						
Font	Message	Age Group 1	Age Group 2	Age Group 3	Age Group 4	Age Group 5
TNR	1	14.34	43.00	32.00	45.67	74.67
TNR	2	16.34	53.67	41.67	42.67	75.70
TNR	3	16.33	45.33	43.33	50.00	74.33
AG	1	15.00	46.34	39.00	44.00	75.00
AG	2	13.67	40.67	32.67	46.66	76.33
AG	3	14.67	42.67	33.00	45.67	75.33

The increase in the spread of data for each group can be attributed to diverse rates of ocular degeneration in different subjects and also to the propensity for ocular disease in the older age groups.

The increase in spread may also be explained by differences in "accommodation". "Accommodation is the term for the eye's ability to focus on an object at varying distances" (8). In the present study, the distance was fixed. Subjects were unable to move the legibility cards, forcing their eyes to "accommodate". "The capacity of the eye to accommodate is gradually lost with increasing age. Human accommodative power declines progressively, beginning in the 2nd decade of life and perhaps earlier, and is completely gone by 75 years of age" (8). Because many of the subjects in the older age groups have probably at least partially lost their power to effectively accommodate, it is logical that the data is much more widely scattered in these age groups.

As mentioned previously, age groups 4 and 5 both contained subjects aged 66-80; these two groups, however, were composed of residents from different states. When the results of the two groups were compared statistically, the difference was found to be significant at α=0.01.(See Appendix F.)

One possible explanation for this difference is that there was an effect of geography or demographic differences were present. Subjects from age group 4 were inhabitants of mid-Michigan in a suburban area. Subjects from age group 5 were inhabitants of Clarinda, Iowa, a rural community where farming is a major source of income.

Subjects from mid-Michigan had primarily been employed in academic or business settings. Group 5, the group from Clarinda, Iowa, primarily consisted of people who earned a living farming. As a result of their occupations, the two groups probably had very different levels of exposure throughout the course of their lives to UV radiation. The "Salisbury Eye Evaluation Project", a study published in the Journal of the American Medical Association (JAMA) (15), suggests that this difference is likely to have a profound impact on the legibility results.

Salisbury researchers used a detailed model to assess the ocular exposure of subjects to UV-B (15) and found a significant association between cortical opacities and average annual UV-B exposure. They also went on to indicate that the risk was a "cumulative dose phenomenon". This offers a second possible explanation for the difference found to exist between groups 4 and 5 in this study. The fifth group, as a result of their occupation, was exposed to more UV-B rays. The probable result is that a higher percentage of the population from group 5 had a significant amount of cortical opacity, or cataracts. This diminished clarity of subjects' lenses translated into higher legibility indices for group 5.

There is a definite need for the ability to objectively measure the legibility of medication labels. If legibility can be effectively measured, then label printing can be improved so that people in the elderly population can use the information provided on labels. This need becomes more and more urgent as the population ages, more prescription drugs are granted OTC status, and self-medication increases in popularity.

Several researchers, including Watanabe (2) and Holt (3), indicate that many of the OTC labels currently on the market are not sufficiently legible to a substantial portion of the elderly population. Even the FDA has suggested that many labels are insufficiently legible. However the most compelling reasons to improve legibility come from the aging users of OTC products. Subjects from every age group in this study indicated that they were aware of deficiencies in the legibility of drug labels. The force of these deficiencies is apparent in the comment of the eighty-year old subject, who indicated, **"I can never read those damn things. So I just take two."**

Things to Consider When Using the Polariscope

Testing older subjects with the polariscope presents problems. Older subjects are more likely to have experienced physiological changes in their eyes that result in lowered light and contrast sensitivity and the inability to accommodate. They are more likely to have ocular diseases such as glaucoma, cataracts and presbyopia. All of these factors may lead to large standard deviations within and between subjects. If this machine is to be used as a performance standard it should be used with younger age groups because they provide more consistent results. (See Appendix C.)

An equation, $y= 0.356x + 5.8325$, has been developed from data collected for this paper which can be used to translate the legibility indices collected from a younger age group into approximate legibility indices for older subjects. It is the equation of the line formed when a bivariate plot of the average age of the age group is plotted against the average legibility index of the same group.(See Figure 10.) This equation is only appropriate with a light level of 25 foot-candles when subjects are viewing cards from a distance of 17.5 inches. It is important that the light level and the distance be consistent because they will have a dramatic impact on the difference in legibility index between the age groups; if the light level increased, the effect of age would not be as significant as it was found to be in this study. Because the power to effectively accommodate is lost over time, the distance from the subject to the card should also be held constant if this equation is to be used. Because group 5 lies outside of the linear relationship established by the other 4 age groups, the equation should not be used to translate scores to ages greater than 73 years.

Summary of Recommendations

1. A study aimed at reducing the variability in legibility indices is strongly advised.
2. A study investigating the effect of different light levels is advised.
3. A study investigating the effect of the distance from the eye to the legibility card is advised.
4. Work to further the use of this instrument as a performance standard is advised. After the variability in subject responses is reduced as much as possible, this information should be used to design an experiment aimed at defining legibility, in terms of a legibility index. This definition should be created with particular concern for the elderly, who have expressed their disgust of the drug labels currently on the market and are a growing percentage of the population.

Appendix A- Consent Form
Label Legibility- Consent to be Tested

Procedure-

Prior to testing, your visual acuity will be measured by reading a card. The lowest line that you can read on the card will determine your visual acuity (20/20, 20/30 etc.). Your color vision will be measured. This information will also be recorded.

You will read a card placed inside the grey box. Look into the box through the round screen on the front. As you look through the screen move the handle on your right 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, push the handle 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. There will be a total of 20 cards for you to read.

You will not be identified by name in any records of this testing; testing is anonymous. This is a test of the printed card, not of your eyes.

Testing will take no more than 20 minutes.

I choose to participate in the label legibility study. ______________________

I decline participation in the study. (Declining to participate will not reflect negatively in any way on subjects). ______________________

If you have any questions or concerns about this study, please feel free to contact Dr. Hugh Lockhart, Michigan State University School of Packaging at 517-355-3604 or Laura Bix at 517-333-9967.

IRB#97-851

Appendix B- Data Recording Sheet

Recording Sheet for Legibility

Measured Visual Acuity: ______________________

Subject #: ______________________

Color Test: ______________________ (Normal Red/Green Color Blind)

19-35 36-50 51-65 66-81

Sex: Female Male

Eyewear: (Circle One) None Single Lens Bifocal Trifocal

Education: 8th Grade High School College Graduate School or Higher

Room Foot-Candles: ______________________

Polariscope Foot-Candles: ______________________

Polariscope Readings

Card #	Reading (Degrees Rotation)	Card #	Reading (Degrees Rotation)

Appendix C-
Descriptive Statistics

Descriptive Statistics

Treatment Combination (Font/Msg/Age)	Mean	Median	Min	Max	Q1	Q3	Standard Deviation
111	16.4	15.7	9.3	23.7	12.3	21.0	4.6
112	21.0	18.2	9.7	52.7	14.8	26.0	9.8
113	25.2	22.2	11.3	43.3	18.6	33.7	9.5
114	34.5	33.5	14.0	59.7	22.2	44.0	12.9
115	48.2	42.2	15.3	90.0	35.6	59.1	19.7
121	16.4	16.0	8.3	24.7	12.3	21.3	4.9
122	21.6	17.2	9.3	63.0	15.1	26.6	11.8
123	25.0	22.0	11.3	53.0	16.9	32.4	10.5
124	33.9	36.7	14.0	56.7	21.9	42.3	11.6
125	45.8	44.0	14.3	90.0	33.6	53.1	18.2
131	16.4	15.5	9.0	25.3	12.0	21.3	4.9
132	20.6	17.7	9.7	55.0	15.0	24.4	9.9
133	25.2	21.2	11.0	54.3	17.7	33.2	11.1
134	34.5	35.0	13.3	63.3	22.0	43.5	12.9
135	45.8	42.3	15.7	90.0	29.6	53.6	19.3
211	15.6	15.0	8.3	23.3	11.8	19.6	4.3
212	19.9	16.7	9.3	55.7	14.1	23.3	10.0
213	23.2	20.8	10.0	49.0	16.7	26.8	10.0
214	32.2	33.5	11.7	55.7	21.8	40.5	11.6
215	44.4	38.7	15.0	90.0	31.8	51.2	19.1
221	15.2	14.7	9.0	22.7	11.4	19.5	4.2
222	19.9	16.8	9.3	50.0	14.8	25.3	9.1
223	22.9	20.3	10.3	43.0	16.3	30.4	8.7
224	31.9	32.2	13.7	60.3	20.2	41.0	11.9
225	45.6	42.0	13.7	90.0	33.9	52.5	18.9
231	15.8	15.0	8.3	23.0	12.2	20.6	4.6
232	19.3	15.5	9.0	51.7	13.8	23.4	9.4
233	23.2	20.8	10.7	43.7	16.4	30.4	9.1
234	31.1	31.5	13.3	59.0	20.8	38.4	10.8
235	44.5	41.3	14.7	90.0	32.6	50.6	18.3

Please refer to Table 1 for legend of treatment combinations.

Appendix D-
Statistical Analysis- All five Groups

Analysis of Variance (Balanced Designs)

Factor	Type	Levels	Values						
Age	fixed	5	1	2	3	4	5		
Font	fixed	2	1	2					
Mess	random	3	1	2	3				
Ns(Age)	random	20	1	2	3	4	5	6	7
			8	9	10	11	12	13	14
			15	16	17	18	19	20	

Analysis of Variance for natural log

Statistical model:

Response = μ + Age_l + $Ns_{j(l)}$ + $Font_k$ + $Mess_l$ + Age*Font + Age*Mess + Ns*Font + Ns*Mess + Font*Mess + Age*Font*Mess + Ns*Font*Mess + $E_{(ijkl)m}$

The restricted form of the mixed model (Mess and Ns are random, the remaining factors are fixed) resulted in the following:

Source	DF	SS	MS	F	P
Age	**4**	**232.5061**	**58.1265**	**22.4**	**4.7*10-13***
Font	**1**	**1.7592**	**1.7592**	**80.4**	**2.9*10-4***
Mess	2	0.0512	0.0256	3.43	0.035
Ns(Age)	**95**	**245.2678**	**2.5818**	**345.42**	**0.000**
Age*Font	4	0.1024	0.0256	1.52	0.217*
Age*Mess	8	0.0916	0.0115	1.53	0.148
Font*Mess	2	0.0270	0.0135	1.57	0.211
Font*Ns(Age)	95	0.8065	0.0085	0.99	0.519
Mess*Ns(Age)	190	1.4201	0.0075	0.55	1.000
Age*Font*Mess	8	0.1123	0.0140	1.63	0.118
Font*Mess*Ns(Age)	190	1.6325	0.0086	0.63	1.000
Error	1200	16.2875	0.0136		
Total	1799	500.0642			

Bolded items were found to be significant at α=0.01.

* Indicates that an approximate F was calculated. Approximate f is calculated in order to obtain the p-value for these effects. Please see the following page for a sample calculation using the effect of age.

Source	Variance component	Error term	Expected Mean Square (using restricted model)
1 Age	*		(12) + 6(9) + 120(6) + 18(4) + 360Q[1]
2 Font	*		(12) + 3(11) + 9(8) + 300(7) + 900Q[2]
3 Mess	0.00003	9	(12) + 6(9) + 600(3)
4 Ns(Age)	0.14302	9	(12) + 6(9) + 18(4)
5 Age*Font	*		(12) + 3(11) + 60(10) + 9(8) + 180Q[5]
6 Age*Mess	0.00003	9	(12) + 6(9) + 120(6)
7 Font*Mess	0.00002	11	(12) + 3(11) + 300(7)
8 Font*Ns(Age)	-0.00001	11	(12) + 3(11) + 9(8)
9 Mess*Ns(Age)	-0.00102	12	(12) + 6(9)
10 Age*Font*Mess	0.00009	11	(12) + 3(11) + 60(10)
11 Font*Mess*Ns(Age)	-0.00166	12	(12) + 3(11)
12 Error	0.01357		(12)

Sample Calculation using the age effect.

Age

Using the Minitab output from the previous page the terms for age are defined as: (12) + 6(9) + 120(6) + 18(4) + 360Q[1]. The objective is to isolate term 360Q[1]. Using the mean square values for terms 1, 9, 4 and 6 an approximate F and the degrees of freedom, p and q, can be calculated. F, p and q are then used in the "F-dist" function of Microsoft Excel to determine the p-value to determine the degree of statistical significance.

F= (Term 1 + Term 9)/(Term 4 + Term 6) = (58.1265 +.0075)/(2.5818 + .0115)=

Approximate F = 22.42

p degrees of freedom =
(Term 1 + Term 9)^2/[((Term 1)^2/df for Term 1) + ((Term 9)^2/df for Term 9)] =
(58.1265 + .0075)^2/[((58.1265)^2/4) + ((.0075)^2/190)] =

p degrees of freedom = 4.00

q degrees of freedom=
(Term 4 + Term 6)^2/[((Term 4)^2/df for Term 4) + (Term 6)^2/df for Term 6)]=
(2.5818 + .0115)^2/[((2.5818)^2/95) + ((.0155)^2/8)]=

q degrees of freedom = 95.83

Using the "F-dist" Function of Microsoft Excel a p-value which can be compared to a predetermined α is obtained.

p value =4.67 * 10^{-13}

The aforementioned p-value indicates a highly significant effect of age.
The effect of age, font and the age*font interaction were calculated in this fashion.

Appendix E- Analysis of 20/20 visual acuities only

Analysis of Variance (Balanced Designs) All five groups included in this analysis using the first 10 subjects with a measured visual acuity of 20/20. Analysis uses natural log as response.

Factor	Type	Levels	Values						
Age	fixed	5	1	2	3	4	5		
Font	fixed	2	1	2					
Mess	random	3	1	2	3				
Ns(Age)	random	10	1	2	3	4	5	6	7
			8	9	10				

Analysis of Variance for natural log

Statistical model:

Response = μ + Age_I + $Ns_{j(I)}$ + $Font_k$ + $Mess_l$ + Age*Font + Age*Mess + Ns*Font + Ns*Mess + Font*Mess + Age*Font*Mess + Ns*Font*Mess + $E_{(ijkl)m}$

The restricted form of the mixed model (Mess and Ns are random, the remaining factors are fixed) resulted in the following:

Source	DF	SS	MS	F	P
Age	4	93.5810	23.3953	*4.00	1.88*10-5
Font	1	0.9146	0.9146	*77.78	2.15*10-5
Mess	2	0.0015	0.0008	0.10	0.905
Ns(Age)	45	115.9639	2.5770	339.72	0.000
Age*Font	4	0.0505	0.0126	*1.53	0.200
Age*Mess	8	0.0692	0.0087	1.14	0.344
Font*Mess	2	0.0113	0.0056	0.51	0.600
Font*Ns(Age)	45	0.2825	0.0063	0.57	0.980
Mess*Ns(Age)	90	0.6827	0.0076	0.48	1.000
Age*Font*Mess	8	0.0750	0.0094	0.85	0.560
Font*Mess*Ns(Age)	90	0.9903	0.0110	0.69	0.984
Error	600	9.5355	0.0159		
Total	899	222.1581			

Bolded items were found to be significant at α=0.01.

* Indicates that an approximate F was calculated. Approximate f is calculated in order to obtain the p-value for these effects. Please see the following page for a sample calculation using the effect of age.

Source	Variance component	Error term	Expected Mean Square (using restricted model)
1 Age		*	(12) + 6(9) + 60(6) + 18(4) + 180Q[1]
2 Ft		*	(12) + 3(11) + 9(8) + 150(7) + 450Q[2]
3 Mess	-0.00002	9	(12) + 6(9) + 300(3)
4 Ns(Age)	0.14274	9	(12) + 6(9) + 18(4)
5 Age*Ft		*	(12) + 3(11) + 30(10) + 9(8) + 90Q[5]
6 Age*Mess	0.00002	9	(12) + 6(9) + 60(6)
7 Ft*Mess	-0.00004	11	(12) + 3(11) + 150(7)
8 Ft*Ns(Age)	-0.00052	11	(12) + 3(11) + 9(8)
9 Mess*Ns(Age)	-0.00138	12	(12) + 6(9)
10 Age*Ft*Mess	-0.00005	11	(12) + 3(11) + 30(10)
11 Ft*Mess*Ns(Age)	-0.00163	12	(12) + 3(11)
12 Error	0.01589		(12)

Sample Calculation using the age effect.

Age
Using the Minitab output from the previous page the terms for age are defined as: (12) + 6(9) + 120(6) +18(4) + 360Q[1]. The objective is to isolate term 360Q[1]. Using the mean square values for terms 1, 9, 4 and 6 an approximate F and the degrees of freedom, p and q, can be calculated. F, p and q are then used in the "F-dist" function of Microsoft Excel to determine the p-value to determine the degree of statistical significance.

F= (Term 1 + Term 9)/(Term 4 + Term 6) = (23.3953 +.0076)/(2.5770 + .0087)=

Approximate F = 9.05

p degrees of freedom =
$(\text{Term } 1 + \text{Term } 9)^2/[((\text{Term } 1)^2/\text{df for Term } 1) + ((\text{Term } 9)^2/\text{df for Term } 9)] =$
$(23.3953 + .0076)^2/[((23.3953)^2/4) + ((.0076)^2/90)] =$

p degrees of freedom = 4.00

q degrees of freedom=
$(\text{Term } 4 + \text{Term } 6)^2/[((\text{Term } 4)^2/\text{df for Term } 4) + (\text{Term } 6)^2/\text{df for Term } 6)]=$
$(2.5770 + .0087)^2/[((2.5770)^2/45) + ((.0087)^2/8)]=$

q degrees of freedom = 45.17

Using the "F-dist" Function of Microsoft Excel a p-value which can be compared to a predetermined α is obtained.

p value = $1.88 * 10^{-5}$

Appendix F- Iowa Versus Michigan- Statistical Analysis

Analysis of Variance (Balanced Designs)

The following analysis includes data from Iowa and Michigan (location 1 is Iowa location 2 is Michigan). Analysis uses natural log as response.

Factor	Type	Levels	Values					
Location	fixed	2	1	2				
Font	fixed	2	1	2				
Mess	random	3	1	2	3			
Ns(Location)	random	20	1	2	3	4	5	6
			7	8	9	10	11	12
			13	14	15	16	17	18
			19	20				

Analysis of Variance for natural log

Statistical model:

Response = μ + Age_l + $Ns_{j(l)}$ + $Font_k$ + $Mess_l$ + Age*Font + Age*Mess + Ns*Font + Ns*Mess + Font*Mess + Age*Font*Mess + Ns*Font*Mess + $E_{(ijkl)m}$

The restricted form of the mixed model (Mess and Ns are random, the remaining factors are fixed) resulted in the following:

Source	DF	SS	MS	F	P
Location	1	17.4197	17.4197	*6.011	0.0189
Font	1	0.5969	0.5969	*21.367	0.0191
Mess	2	0.0417	0.0208	2.12	0.127
Ns(Location)	38	110.1137	2.8977	295.18	0.000
Location*Font	1	0.0530	0.0530	*1.44	0.352
Location*Mess	2	0.0033	0.0017	0.17	0.844
Font*Mess	2	0.0420	0.0210	2.68	0.075
Font*Ns(Location)	38	0.2790	0.0073	0.94	0.576
Mess*Ns(Location)	76	0.7461	0.0098	0.64	0.992
Location*Font*Mess	2	0.0696	0.0348	4.45	0.015
Font*Mess*Ns(Location)	76	0.5945	0.0078	0.51	1.000
Error	480	7.3901	0.0154		
Total	719	137.3494			

Bolded items were found to be significant at α=0.01.

* Indicates that an approximate F was calculated. Approximate f is calculated in order to obtain the p-value for these effects. Please see the following page for a sample calculation using the effect of age.

Source	Variance component	Error term	Expected Mean Square (using restricted model)
1 Loc	*		(12) + 6(9) + 120(6) + 18(4) + 360Q[1]
2 Font	*		(12) + 3(11) + 9(8) + 120(7) + 360Q[2]
3 Mess	0.00005	9	(12) + 6(9) + 240(3)
4 Ns(Loc)	0.16044	9	(12) + 6(9) + 18(4)
5 Loc*Font	*		(12) + 3(11) + 60(10) + 9(8) + 180Q[5]
6 Loc*Mess	-0.00007	9	(12) + 6(9) + 120(6)
7 Font*Mess	0.00011	11	(12) + 3(11) + 120(7)
8 Font*Ns(Loc)	-0.00005	11	(12) + 3(11) + 9(8)
9 Mess*Ns(Loc)	-0.00093	12	(12) + 6(9)
10 Loc*Font*Mess	0.00045	11	(12) + 3(11) + 60(10)
11 Font*Mess*Ns(Loc)	-0.00252	12	(12) + 3(11)
12 Error	0.01540		(12)

Sample Calculation using the location effect.

Effect of location
Using the Minitab output from the previous page the terms for location are defined as: (12) + 6(9) + 120(6) +18(4) + 360Q[1]. The objective is to isolate term 360Q[1]. Using the mean square values for terms 1, 9, 4 and 6 an approximate F and the degrees of freedom, p and q, can be calculated. F, p and q are then used in the "F-dist" function of Microsoft Excel to determine the p-value to determine the degree of statistical significance.

F= (Term 1 + Term 9)/(Term 4 + Term 6) = (17.4197+.0098)/(.0017 + 2.8978)=

Approximate F = 6.01

p degrees of freedom =
$(\text{Term } 1 + \text{Term } 9)^2/[((\text{Term } 1)^2/\text{df for Term } 1) + ((\text{Term } 9)^2/\text{df for Term } 9)] =$
$(17.4197 + .0098)^2/[((17.4197)^2/1) + ((.0098)^2/76)] =$

p degrees of freedom = 1.00

q degrees of freedom=
$(\text{Term } 4 + \text{Term } 6)^2/[((\text{Term } 4)^2/\text{df for Term } 4) + (\text{Term } 6)^2/\text{df for Term } 6)]=$
$(2.8978 + ..0017)^2/[((2.8978)^2/38) + ((.0017)^2/2)]=$

q degrees of freedom = 38.04

Using the "F-dist" Function of Microsoft Excel a p-value which can be compared to a predetermined α is obtained.

p value = .018922

Bibliography

1. Food and Drug Administration, Federal Register, "Proposed Rule on Labeling of OTC Drug Products", February 27, 1997. Pp. 9024-9062.

2. Watanabe, Ronald K. O.D. "The Ability of the Geriatric Population to Read Labels on Over-the-Counter Medication Containers", Journal of the American Optometric Association, Vol 65, No. 1, January 1994. Pp.32-37.

3. Holt, Gary; Hollon, James; Hughes, Scott; Coyle, Rebecca, "OTC Labels: Can Consumers Read and Understand Them?", American Pharmacy, Vol NS30, No.11, November, 1990. Pp.51-54.

4. Gitau, Mutane Wa, Legibility of Child Resistant Bottle Caps by Two Methods of Measurement, Graduate Thesis, Michigan State University School of Packaging, 1988.

5. Lockhart, Hugh, PhD. Packaging 480. Spring 1997.

6. Lockhart, Hugh, PhD.; Bix, Laura from unpublished file report. "Comment to FDA on Proposed Rule". July 1997.

7. Pietrowski, James. Development of a Methodology to Quantify Package and Label Legibility, Graduate Thesis, Michigan State University School of Packaging, 1993.

8. Kelly, Maryann. Visual Impairment in the Elderly and its Impact on Their Daily Lives, PhD Dissertation, Texas Woman's University College of Nursing, 1993.

9. Marieb, Elaine N. Human Anatomy and Physiology, Benjamin/Cummings Publishing Company, Inc., 1992.

10. Holt, Gary; Hall, Edwin, "The Self-Care Movement", Handbook of Nonprescription Drugs, 9th ed., American Pharmaceutical Association,1990. Pp. 1-9.

11. Caranasos, George, MD; Stewart, Ronald, MS; Cluff, Leighton,MD, "Drug-Induced Illness Leading to Hospitalization", Journal of the American Medical Association, Vol 228, No. 6, May 6, 1974. Pp. 713-717.

12. Dow Corning Ophthalmics, Near Point Visual Acuity Card Instructions, Norfolk, VA, 1981.

13. Montgomery, Douglas, C., Design and Analysis of Experiments, John Wiley and Sons, 1997. Pp. 487.

14. Gilliland, D; Bix, L.; Chen, B; Sung, H. from unpublished file, "Using the Polariscope as a Measure of Legibility". December, 1997.

15. West, S; Duncan, D; Munoz, B; Rubin, G; Fried, L; Bandeen-Roche, K; Schein, O "Sunlight Exposure and Risk of Lens Opacities in a Population-Based Study", Journal of the American Medical Association, Vol 280, No.8, August 26, 1998. Pp.714-718.