IMPROVING INTERACTIONS BETWEEN SELF-MEDICATING CONSUMERS AND OVER-THE-COUNTER PACKAGING WITH FRONT-OF-PACK AND PERSONALIZED LABELING AS STRATEGIES

By

Lanqing Liu

A DISSERTATION

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

Packaging – Doctor of Philosophy

2022

ABSTRACT

IMPROVING INTERACTIONS BETWEEN SELF-MEDICATING CONSUMERS AND OVER-THE-COUNTER PACKAGING WITH FRONT-OF-PACK AND PERSONALIZED LABELING AS STRATEGIES

By

Lanqing Liu

Interactions between self-medicating consumers and labeling of Over-the-Counter medications (OTC) influence quality of information processing and hence appropriateness of medication decisions. Our previous work on human-package interaction yielded evidence that early stages of processing important regulatory information were necessary to improve OTC packaging labeling and human-package interactions, and thus to inform appropriate decisions.

Under the framework of Human-Package Interaction model (H-PIM) and the types of directiveness of label designs, we proposed two novel labeling strategies: Front-of-Pack (FOP) labeling and personalized FOP labeling. The FOP strategy utilized the concepts of front-of-pack, boxing, grouping, and highlighting (HL), whereas the personalized FOP strategy further combined the concepts of the FOP labeling with augmented user interface and decision-support signals to assist the decision-making process for enhancing human-package interactions.

To quantitatively investigate the effectiveness of those FOP labeling strategies, we firstly conducted a *change detection test* to evaluate the impact of the FOP labeling strategy on consumers' attention to critical drug information. Additionally, we then developed an *absolute judgement test* to evaluate the effectiveness of the personalized FOP labeling strategy for assisting decision-making to benchmark the potential benefits of this strategy.

The *change detection test* results indicated that the use of HL was effective and efficient to garner attentions. Specifically, the presence of HL increased change detection accuracy

(HL:ME=0.932, SE=0.008; not HL:ME=0.770, SE=0.019; p<0.001)) and shortened the time to correctly detect changes. (HL: ME=3.790, SE=0.200; not HL:ME=5.073, SE=0.268; p<0.001). However, no evidence was found to suggest that the use of FOP labels enhanced the change detection accuracy. Moreover, the presence of FOP labels could prolong the time consumers used to correctly detect changes on the OTC packages than the standard labels. (FOP: ME=4.542, SE=0.238; standard: ME=4.233, SE=0.225; p=<0.001) These results may be caused by factors such as the FOP label location and unbalanced experimental design. Further studies are needed to gain more knowledge of this strategy.

The *absolute judgement test* results supported the effectiveness and efficiency of the personalized FOP strategy on improving decision appropriateness. When introduced and educated with the personalized labeling concept, participants made decisions significantly more accurate (personalized FOP: ME=0.977, SE=0.007; standard: ME=0.933, SE=0.017; p=0.002) and faster (personalized FOP: ME=9.584, SE=0.854; standard: ME=19.052, SE=2.322; p<0.001) with the presence of personalized FOP labels compared to the presence of standard ones.

To conclude, this dissertation extends FOP strategies from non-directive labels to personalized labels. The personalized labeling could act as an important role in improving the interactions between consumers and OTCs. Future studies are needed to gain more knowledge on effectively presenting the strategies as well as applying them to a broader range of package types, populations, environments, and etc.

To Grandpa,
I wish I could help your package rage on those medicines.

ACKNOWLEDGEMENTS

To Dr. Laura Bix, I harbor great gratitude to you for your mentorship. Your guidance helped me grow as a trained researcher, a scholar, and a person. What I learned from you will serve for my lifetime. To Dr. Mark Becker, I enjoy discussing research with you. The innovative ways you looked at problems always inspired me with many "aha" moments. To Dr. Rafael Auras, I am grateful for your critical thinking and academic attitude, which helped me lay a solid foundation as a qualified researcher. To Dr. Susan Selke, I am honored to have you in the committee. This work cannot be possible without your expertise and invaluable advises.

To Dr. Andrew Dennhardt and Dr. Chitrak Andrew from CSTAT, I would like to thank for your expertise in statistics and support for the data analysis of the change detection study. To Sukhdeep Singh from the CANR Statistic Consulting Center, your patient explanations for my statistical questions were very helpful for my data analysis of the absolute judgement test.

To my Packaging HUB colleagues: Dr. Jiyon Lee, Dr. Tony Trier, Shiva Esfahanian, Dr. Alyssa Harben, Dr. Cory Wilson, Krishnaa Venkatesan, Goheth Motamarri, Eric Estrada, Eric Brunk, Dr. Paula Perez, Donnie Kirkland, Dr. Hao Liu, Dr. Dochan Seo, Dr. Raghav Sundar and Dr. Javier de la Fuente. Thanks for your encouragements, your critics, as well as your supports on pilot testing and data collection. The friendship with you is a great treasure of my life.

To my wife, Dr. Ying Cheng, thanks for your companionship and support. Your love makes everything better during this pandemic. To my parents, thanks for your expertise in clinical medicine as well as your great love. I owe a lot to you. Lastly, to my little mover Clara, with love from your dad. You are so lovely, but to be honest, your current behavior further affirms that research on child-resistant package is a must.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

ADR	Adverse Drug Reaction
AR	Augmented Reality
CFR	Code of Federal Regulation
DFL	Drug Facts Label
FDA	US Food and Drug Administration
FFDCA	Federal Food Drug and Cosmetic Act
FOP	Front-of-Pack
GDA	Guideline Daily Amounts
H-PIM	Human-Package Interaction Model
HL	Highlighting
HMD	Head-Mounted Display
OTC	Over the Counter
PDP	Principal Display Panel
US	United States
VR	Virtual Reality

Chapter 1 Introduction and Review of Literatures

1.1 Introduction and Literature Review

1.1.1 Over-the-Counter Medications and Self-medicating Patients in the U.S.

Over-the-Counter medications refer to the medications that patients can buy without a prescription. In US markets, OTCs are regulated by the US Food and Drug Administration (FDA) through the Federal Food Drug and Cosmetic Act (FFDCA). (*CFR - Code of Federal Regulations Title 21*) By definition, OTCs are considered to be safe and effective for use by the general public without seeking treatment from a health professional. Because of the benefits such as cost-saving, convenience, easy-access and flexibility, OTCs play an increasingly vital role in America's health care system. (*OTC Retail Sales 1964-2019*, n.d.) There are over 300,000 marketed OTC drug products in the U.S., and the Food Drug Administration (FDA) reviews the active ingredients and the labeling of over 90 therapeutic classes/categories of drugs, such as analgesics, antacids. (FDA, 2020)

Self-medication is defined as the act taking of drugs (including both prescriptions and OTCs), herbs, and home remedies, on one's own without consultation of medical professionals. (Bennadi, 2013; *Guidance for Industry Labeling OTC Human Drug Products*, 2009; Shehnaz et al., 2014; Zhao & Ma, 2016) The process of self-medicating involves recognition of symptoms, selections of therapies (including medicines), and interpretation of (and appropriate action on) dosage and schedule. Self-medication is one of the essential components of self-care, a broader term, which includes all health-related decision-making by individuals and family members. (Mahapatra, 2017)

1.1.2 Potential Risks of Self-Medication with OTCs

Despite its popularity, self-medication with OTCs comes with risk. Simple, and routine decisions about OTCs can have negative consequences. These consequences are more prevalent in vulnerable populations (e.g., aging, those with poor literacy, non-native speakers) as well as those engaged in complex drug regimens. Negative consequences, or adverse drug reactions (ADRs) can be the result of drug-drug interactions or drug diagnosis interactions; these are of particular concern for people with multiple comorbidities, where drugs can be contraindicated with existing underlying conditions or other treatments. While labeling information is critical for all self-medicating patients, it is of particular importance to populations more likely to suffer an adverse drug reaction (ADRs).

ADRs can be defined as "an appreciably harmful and unpleasant reaction resulting from an intervention related to the use of a medicinal product." (Aronson & Ferner, 2005) Traditionally, ADRs have been classified into two types: (1) Type A reactions refer to as augmented reactions which are "dose-dependent" and predictable based on the pharmacology of the drug. (2) Type B reactions refer to bizarre reactions which are not predictable based on the pharmacology. (Coleman & Pontefract, 2016)

Unlike prescriptions, OTCs lack any mandatory supervision under a learned intermediary during selecting and dosing, which makes the labeling an important intervention for improving consumer the use of medication and the understanding of safety information, and hence helpful in preventing the potential occurrence of Type A ADRs, such as overdosing, drug-drug or drug-diagnosis interactions. (Schmiedl et al., 2014)

1.1.3 Over-the-Counter Medication Packaging Labeling

Labeling has been identified as a common strategy for delivering important information. Consumers have indicated OTC labeling as a preferable source for information when making medical decisions (Westerlund et al., 2017). When self-medicating, OTC labels provide patients specific information important for a medications safe and effective use, including active ingredients, directions, warnings, and dosage information intended to enable them to select and administer a given medicine appropriately. (Tong et al., 2017)

Recognizing the important role of labels to these products, the US Food and Drug Administration (FDA) has long-regulated OTC labeling with specific requirements for information content within Title 21 of the Code of Federal Regulations Part 201 Subpart C (*CFR - Code of Federal Regulations Title 21*, n.d.) and for the formatting of the same in Subpart D; additionally, they have provided a Guidance for Industry intended to assist with the development of labeling for OTC human drug products. (*Guidance for Industry Labeling OTC Human Drug Products*, 2009)

There are two major components that comprise regulated information for OTCs: (1) the Principal Display Panel (PDP) (21 CFR 201.66), defined as "the part of a label that is most likely to be displayed, presented, shown, or examined under customary conditions of display for retail sale," and (2) the Drug Facts Labeling (DFL) (21 CFR 201.66) which includes "the active ingredients and their purpose, the product's uses, warnings, directions, other information, and inactive ingredients." Our previous work, (Liu, 2016), provides a summary of Title 21 CFR 201 pa.66 parts C (information content) and part D (formatting) required for the DFL. It is worth noting that DFLs, which contain the majority of dictated regulatory information, are intended to "make it easier for consumers to read and understand OTC drug product labeling and use OTC

drug products safely and effectively." (*Guidance for Industry Labeling OTC Human Drug Products*, 2009)

1.1.4 Consumer Interactions with OTC packages and Decision Making

To make an appropriate decision, it is essential for patients to utilize information available in the DFL that is germane to their personal needs and the context requiring the drug (e.g. their individual health history, current medications, present condition, needs and state of health at the time of decision making). For example, based on in-store shopper observations and laboratory-based simulated OTC shopping tasks among older adults, Holden et al. (2019) found that people searched for medication adverse effects and safety information; a key finding of the research team was that participants primarily relied on packaging during decision making (in lieu of pharmacy staff). Viewed through an information processing frame (DeJoy, 1991), to be effective, relevant labeling information must be noticed, carefully read, and thoroughly understood. Specifically, the DFL contains the necessary information to help patients engaged in self-medication understand the active ingredients present in products that they are considering, as well as important warnings, which may include drug/drug and drug/diagnosis contraindications, as well as directions for appropriate use.

However, available research suggests that a lack of engagement with all types of information on OTC labels, particularly information contained on the DFL, is endemic and problematic. Available work suggests that consumers fail to attend to active ingredient or the related warning information on OTC labeling during drug selection. King et al. (2011) found that only 41% of their participants indicated that they always look at active ingredient information when purchasing an OTC drug. Similarly, another survey indicated that 78% of respondents use symptom relief in guiding purchase decisions; 54% use brand name; 47% look for sale products;

with no mention of information about the active ingredient, or disease or drug contraindications guiding purchase decisions. (Aker et al., 2014)

1.1.5 Human-package Interaction Model (H-PIM)

Since Card's seminal work on the Human Processor Model (S. Card et al., 1986 & 2018), Various models have been proposed to organize how people process information related to external stimuli in order to perform tasks and make decisions. Advancing this line of theoretical work, researchers have applied various iterations of these models in an attempt to organize and understand how people process labeling information present on packaging to make decisions related to medical products as they consider and use them.(Shaver & Wogalter, 2003; Berman, 2004; King et al., 2011; Tong et al., 2014; Trivedi et al., 2014; Laura Bix et al., 2015; Liu, 2016; L. Bix et al., 2016; Esfahanian & Link to external site, 2020) The following section will explain the theoretical background which undergirds and organizes the study proposed herein.

The Human Processor Model postulates that humans employ their perceptual, cognitive and motor systems to process information and act upon it. The perceptual system handles sensory stimulus from outside world (i.e., the five senses), and motor system controls actions to accomplish a task based on the information. The cognitive system supplies processing to connect the perceptual system (input) and motor system (output).

The Cyclic Interaction Model (Monk, 1999) further specified an input-output process for people making decisions. Specifically, this model proposes a cyclic information flow with six stages of human-product interactions: exposure, perception, encodation, comprehension, execution, and action. During an interaction, the information must be *exposed* to target users to make it perceptible via the senses (stage 2); specific to OTC products, *perception* would generally occur through the use of a label that is perceived visually. Perceived information is then *encoded* (stage 3) into an internal representation capable of being recognized and assigned

meaning or thought by users. In the context of processing OTC labeling information, this stage involves the visual image landing on the retina in the eye for translation into an electrical impulse that can be interpreted by the cognitive system. There are limited resources available for processing by the cognitive system, so depending on the context surrounding the user (distractions, or devotion of resources to processing of other tasks), it is conceivable that although the visual is being brought into the eye, it may not necessarily be encoded for further processing. After the signal is encoded the stage of *comprehension* begins; if the reading level of the text is beyond the capability of the viewer or a symbol is confusing, the viewer may fail to interpret it correctly. After reconciling information that they have gathered about the products during comprehension with previous knowledge and experiences (e.g., an existing medical condition or a drug that they are currently taking that are contraindicated with the product under consideration), the viewer is moved to the fifth stage of the model, *execution*, by utilizing the motor systems to take action based on their assessment of the information. This action changes the state of things accomplishing the task (e.g., product selection, appropriate dosing amount, etc.) and starts the cycle anew with the next text. This cycle repeats until a user's goal (a series of tasks) is achieved.

The Human Packaging Interaction Model H-PIM (de la Fuente et al., 2015), combined and adapted the work of Card (Card et al., 1986) and Monk (Monk, 1999) and Shackel & Richardson (the Usability Theory) (Shackel & Richardson, 1991). Specifically, the HPiM suggests that each of the five stages of the aforementioned information processing model is impacted by the four inputs of the Usability theory. Specifically, the Usability Theory postulates four principal components (or inputs) encapsulate a human-technology interaction, namely *user*, *task*, *tool*, and *environment*. The user input refers to the characteristics of the person, including

their perceptual, cognitive, and physical capabilities, habits, behaviors, abilities, beliefs, previous experiences, etc. The task input involves a single step toward a goal that users seek to achieve (e.g. selection of a product that is safe for them to use). In Shackle's terms, the tool input represents the object (e.g., technology, product, machine) to interact with (in our case an OTC label design). The context input includes the physical, social, and cultural environment, including things such as lighting, seating, distractions, and conventions related to appropriateness etc. The functionality of the system (a person's ability to navigate information processing) depends on the dynamic interplay between the four components. For example, successful system design for tools of any type 1) allows adaptations to different users, tasks, and environment, 2) are easy to use, and 3) help users accomplish tasks effectively.

Although the Usability Theory was originally developed to evaluate the usability of a technology tool, this theory can also be applied to understand how OTCs packages (i.e., the tool) can help people with various abilities, habits, behaviors, and beliefs (i.e., user- for instance, those with poor health literacy) to make appropriate medical decisions (i.e., task) in retail pharmacies (i.e., context). Putting these concepts to practice challenges designers to develop medical packaging with maximum usability. That is, a convenient solution (i.e., ease-of-use) capable of being navigated by people of diverse abilities and backgrounds (i.e., flexibility) to make more accurate health decisions (i.e., effectiveness).

Figure 1.1 provides a visual of the H-PIM model (de la Fuente et al., 2015) and depicts that users' actions are undergirded by context when a user/consumer performs a task(s) with a packaged product. As with Monk's original proposal, the interaction is a cyclic process, with the action potentially producing an effect that resets the state of things, beginning the information processing portion anew as the user begins to accomplish subsequent tasks. It provides a

comprehensive framework for analyzing behaviors of self-medicating consumers when they are selecting or using OTC medications and can also be used to organize and coordinate considerations related to experimental design related to package evaluation.

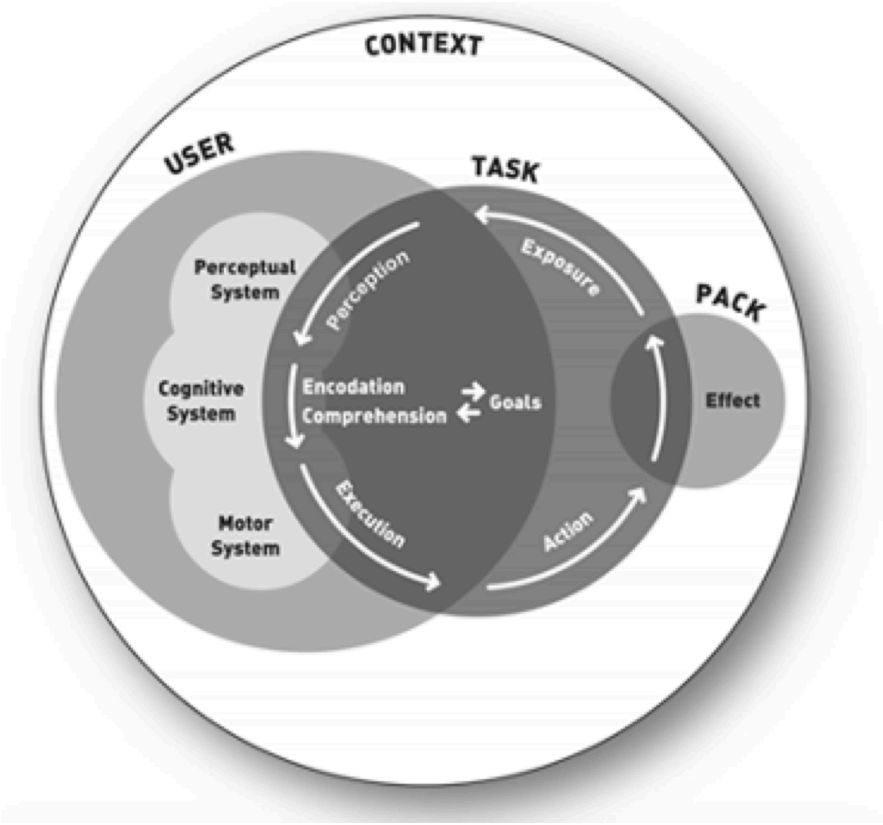


Figure 1.1 Human-Package Interaction Model by Javier de la Fuente. Reprinted with permission, dela Fuente CJ (2013) Usability of tabs in semi-rigid packaging (Doctoral Dissertation). School of Packaging, Michigan State University.

In this study, we seek to optimize packaging design strategies to address information processing challenges during the exposure, perception, and comprehension stages of decision-making related to OTC drugs. Our ultimate goal is improving self-medicating patients' engagement with, and understanding of, information critical to the safe and effective use of OTC products (active ingredients and drug/drug and drug/diagnosis warning information) and, ultimately inspiring appropriate decision making. The following section explains the challenges of making appropriate medical decisions at each of the information processing stage.

Table 1.1 Stages of Human-Package Interaction Model

Stage	Description
Exposure	User is exposed to necessary information.
Perception	Information is perceived by user's sensory system.
Encodation	Perceived information is transformed into an internal signal that can be further processed by cognitive systems.
Comprehension	User recognizes and assigns meaning to the encoded information.
Execution	Thought is translated into actions.

1.1.6 Exposure of Information and Front-of-Pack Labeling

In accordance with the reviewed models (S. Card et al., 1986; de la Fuente et al., 2015; Monk, 1999), information processing occurs in a serialized fashion; in order for a user to be successful processing the information being provided, they must proceed through the steps in order (i.e., for information to be effective, they must first be exposed to it, then perceive it, etc. through to action See Table 1.1). Under this paradigm, if a consumer fails to be exposed to, perceive, or encode information necessary for the safe and effective use of an OTC product, they are unable to further process it (comprehension, and ultimately action). Research suggests that early-stage processing (Stages 1-3) is problematic when people interact with OTC products. For example, in a study conducted among adults 65 years or older by (Liu, 2016), 50% of the participants focused solely on the PDP information never referring to the comprehensive information present in the DFL (Figure 1.2), suggesting participants failed to be exposed (Stage 1- See Table 1.1) to the required regulatory information that can be critical to making informed health decisions for some consumers. As exposure is prerequisite for further processing, developing a packaging strategy that enhances early-stage processing of information that is important to the safe and effective use of OTCs is desirable.

A review of the literature focused on Front-of-Pack (FOP) labeling related to food packaging suggested that the use of truncated information nutrition information on the PDP of

food packaging results in enhanced attention to, and comprehension of, nutrition information compared with traditional formats of nutrition information. (Cowburn & Stockley, 2005; Hawley & Leasure, 2012; Hersey et al., 2013; Ikonen et al., 2020)

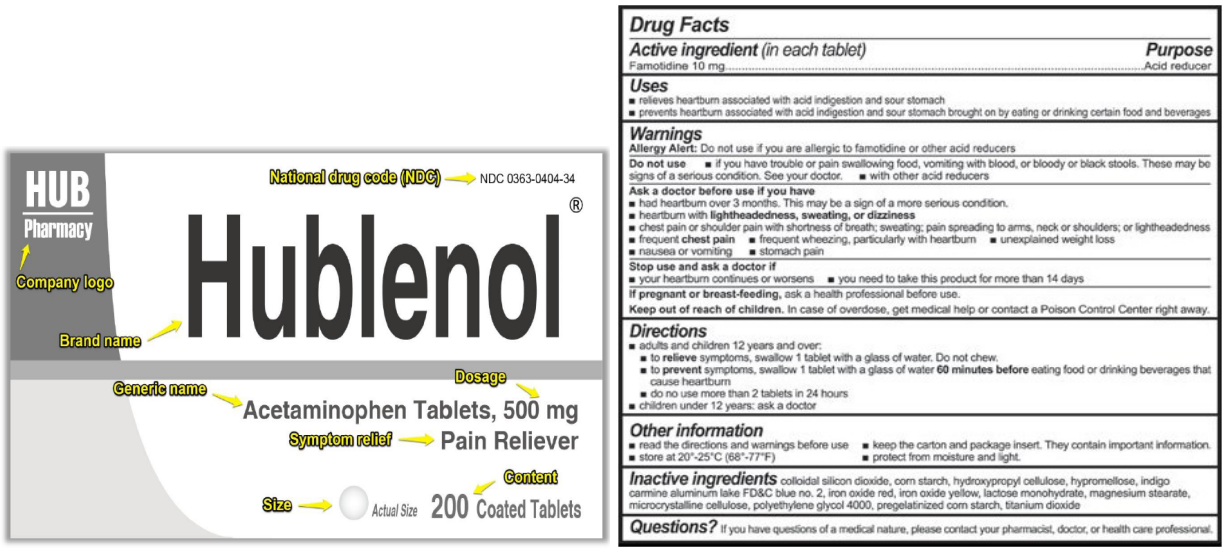


Figure 1.2 Principal display panel (PDP) and Drug facts label (DFL).

In food products the FOP strategy places key information regarding nutrients associated with diseases (e.g., sodium, sugar, fat, and saturated fat) on the front of the package, or Principal Display Panel (PDP) with the motivation of finding ways to induce consumers to make healthier diet choices (Nijman et al., 2007). By definition, the PDP is the face that is customarily displayed at retail and has been noted to be more commonly viewed by at risk consumers interacting with OTC (i.e., information displayed there is more likely to be exposed). As a result, *we postulated that placing critical information from the Drug Facts Label (DFL) onto the PDP in the form of a front of pack warning label (FOP) would result in enhancements in early-stage processing when compared with OTC labels fashioned on existing commercial and regulatory standards*.

This hypothesis is supported by our review of the literature regarding FOPs used in food products, which suggests that the approach attracts attention from consumers more readily (L. Bix et al., 2016) as well as simplifies product comparisons (Hersey et al., 2013) and ultimately,

leads to better decision-making. For example, Roseman et al. (2018) reported that participants selected more nutrient-dense snacks when the product featured an FOP label. Since its introduction, research investigating FOP labeling strategies has increased considerably, providing evidence that various FOP schemes help people accurately interpret product information and make healthy choices (Feteira-Santos et al., 2020).

FOPs that employ a qualitative assessment related to relative health values (e.g., color, i.e., traffic light symbols; red high dietary levels yellow for moderate levels and green for low levels) have been found as more likely to facilitate healthier selections than FOPS merely displaying numeric information about a product (e.g., Guideline Daily Amounts scores). (Hersey et al., 2013) From the standpoint of information processing, the benefit could be associated with benefits to early-stage processing (enhanced attention due to the use of color), or lower cognitive loads in later stages of processing; that is, that a reduced cognitive burden associated with the interpretation of the information assessing product healthfulness, eases decision-making. In other words, by "doing the work" regarding the evaluation of healthfulness, the FOP scheme reduces the chance for people to misinterpret information on packages, easing late-stage processing (See Table 1.1- Steps 4-6). This may be an especially helpful approach for those who have literacy issues, read English as a second language, or have complex medical regimens to consider. Consistent with this idea, research into FOP labeling strategies in food products suggests that people prefer simple, straightforward designs to complex labeling strategies. Feunekes et al. (2008) tested the effectiveness of eight FOP nutrition labeling schemes that differed in complexity across four countries in Europe. The labeling formats varied from simple (e.g., healthier choice ticker, health protection factor, stars, and smiles) (Figure 1.3 a to d) to the more detailed, comprehensive formats (e.g., multiple traffic light, wheel of health, multiple choice

ticker and Guideline Daily Amounts (GDA) scores). (Figure 1.3 e to f)) The simple formats summarized the whole nutritional profile and provided an overall interpretation of the healthiness of the product, while others provided explicit information on key nutrients. Researchers concluded that the simplified FOPs helped people make healthier choices in a retail environment compared to more complex FOP formats because the former format allowed for shorter processing time and was easier to understand (late-stage processing see Table 1.1- Stage 4).

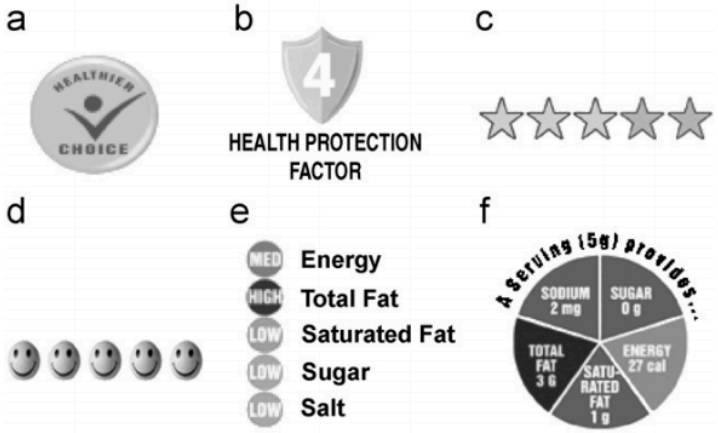


Figure 1.3 The Nutrition Labelling Formats used in Study 1: (a) Healthier Choice Tick; (b) Health Protection Factor; (c) Stars; (d) Smileys; (e) Multiple Traffic Lght; (f) Wheel of Health. (Feunekes et al., 2008)

However, one drawback of FOP labeling that offers an overall evaluation of a food product, is that although this strategy can enhance perceived healthfulness of healthful products, it cannot reduce perceived healthfulness of unhealthful products (Cabrera et al., 2017). The concern is especially pertinent to OTCs given the potential for ADRs. To address this concern, warnings have been recently proposed as a new type of FOP labeling design to flag high content related to nutrients associated with disease (e.g., fat), dangerous practices, or risks (Cabrera et al., 2017; Gawasane et al., 2012). Putting warnings on the FOP can increase the visibility of the information given research suggesting that many people do not turn away from the product's PDP during decision making (Liu, 2016). As a support for the effectiveness of warnings in the

form of an FOP, Arrúa et al. and colleagues (2017) found that using the strategy of prioritizing nutrients of concern (warnings) resulted in equal performance to the traffic-light FOP scheme to help people identify the most healthful product. Moreover, warnings outperformed FOPs featuring traffic-light and Guideline Daily Amounts to help people identify the least healthful products.

We envision that by combining emerging technologies (e.g., artificial intelligence and augmented reality) in light of the promising research which support the use of FOP labels for food products, OTC labels could be reinvented in ways that significantly enhance decision making. Herein, we propose that completely rethinking OTC warning label design has the potential to enhance attention (early-stage processing) and comprehension (late-stage processing).

1.1.7 Personalized Labeling and Augmented Reality

The use of technology, such as virtual reality (VR) and augmented reality (AR), in patient care and medical practice has become increasingly prominent in recent years (Parekh et al., 2020; Takemoto et al., 2019) VR seeks to create an artificial environment in which a person can experience and explore interactively (Höllerer & Feiner, 2004). Immersion into the virtual world is often fulfilled by taking over a person's entire vision using a head-mounted display (HMD) (Sutherland et al., 2019). Different from VR, AR attempts to map virtual objects or annotations onto the real world (Bin et al., 2020; Parekh et al., 2020). The implementation of AR, at minimum, requires 1) positional tracking of the user's eyes or head to determine the image and perspective to display and 2) visualizing virtual objects from the user's perspective (Sutherland et al., 2019). To achieve the visualization of virtual objects in the real world, an AR system can rely on expensive devices such as HMDs or see-through glasses or the use of simple, handheld displays, such as smartphones (Moro et al., 2017; Sutherland et al., 2019). The latter enables AR

(vs. VR) to be a more feasible system for patients' medical self-care because of its mobility, portability, and the increasing prevalence of these devices. As such, this study focuses on exploring the concept of using AR to help people make medical decisions.

Research on the application of AR in medicine has been largely focused on its utility for healthcare professionals, such as pharmacy education and surgery (Moro et al., 2017). For example, Tran et al. (2011) developed an AR system that superimposed three-dimensional virtual presentations of osseous structures and soft tissue on patients' body, preventing surgeons from penetrating into high-risk areas during an oral surgery. Despite the growing body of research on the application of AR for healthcare professionals, only a few studies have examined how AR systems can facilitate patients' self-care (Takemoto et al., 2019). In a study conducted by Diodati et al. (2015), researchers created a mobile AR application that could capture labeling information from medical packages and combine this information and personal health records to help healthcare professionals evaluate patients' self-care quality. This study represents an important step in realizing the potential AR has to provide personalized information based on patients' health conditions. However, the AR application developed by Diodati et al. (2015) still relies on healthcare professionals to make health decisions for patients. Yet, recent events, including: OTC Monograph reform, which opens the door for more timely and flexible OTC regulation; the restoration of eligibility of OTCs under tax-preferred HAS and FSA accounts; as well as increased trepidation to seek formal healthcare and escalating financial concerns during the pandemic have led experts to predict robust, continued growth of this sector suggest that self-medicating is an important (and growing) trend. (Melville, 2021) In fact, for some patients, self-care may be the only option (Takemoto et al., 2019). To take advantage of AR's potential to fulfill personalized medication and enhance patients' self-care, our eventual goal is the creation

an AR smart phone application that utilizes AR to layer a virtual FOP label over an OTC product as a consumer considers it for their use, returning an answer of its appropriateness based on their own health history and current medication history.

1.2 Research Goals

In support of the overarching goal to implement interactions between self-medicating consumers and OTC packaging we proposed the following Aims using strategies which leveraged from the FOP labeling strategy to the personalized labeling strategy with augmented user interface to encourage attention garnering and appropriate decisions.

In *Aim 1*, the non-directive FOP labeling strategy was introduced and its effectiveness on attentive behaviors to critical drug information was investigated. This provided pilot data regarding attention as a function of location (PDP vs DFL). In *Aim 2*, we proposed the framework of the personalized labeling strategy and focused on the development of its user interface with augmented reality technology. In *Aim 3*, we objectively investigated the effectiveness of the personalized labeling concept on assisting people's decision-making process on OTC appropriateness.

Chapter 2 Front-of-Pack Labeling Strategy as a User-centered Approach for Over-the-Counter Medications (Aim 1)

2.1 Objectives

The goal of the Aim 1 was to objectively evaluate the efficacy of an FOP strategy for OTC labels. Specifically, this study provided pilot data regarding attention as a function of location (PDP vs DFL) using a change detection method. Critical changes which occurred in information required for the safe and effective use of the OTCs being tested were evaluated for a location effect for different combinations of label layout (FOP vs Standard) and highlighting (Highlighting vs Not highlighting).

Noting that the critical drug information could change both within FOP box and outside of FOP box in treatments which employed the FOP strategy, our specific research questions for the Aim 1, therefore, were:

- **RQ1:** when changes were outside of the FOP, what were the differences in the ability to garner participants' attention in FOP treatments compared to performance when the standard label was tested. Both accuracy and time to correct response served as dependent variables in the evaluation.
- **RQ2:** What differences in participants' performance (both accuracy and time to correct selection) existed when performance on changes occurring inside the FOP box, were compared with the same type of information that changed outside of the FOP box?

2.2 Method

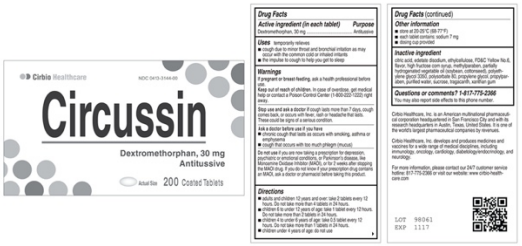
This change detection test was built and run using E-prime 3.0. (*Tools Psychology Software, Pennsylvania, USA*). The test was designed following Rensink's change detection timings and set up. (*Rensink, 2002*) This study was reviewed and approved by MSU IRB under the number i053638.

Stimulus images were designed in greyscale with the resolution of 1920x1080 using Adobe Illustrator CS6. Stimulus were built using three common active ingredients, each of them represented a different category of drug: ibuprofen (IBU) - a pain reliever, dextromethorphan (DEX) -a cough and cold treatment, and ranitidine (RAN) – an acid reducer. For each active ingredient, we developed a mock brand: Hexidvil for ibuprofen, Circussin for dextromethorphan and Recantac for ranitidine. (Table 2.1)

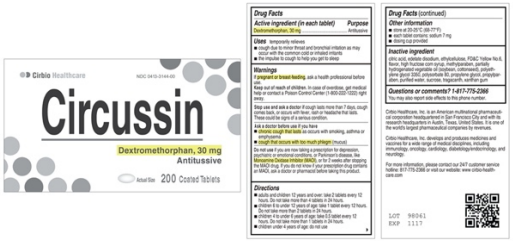
Table 2.1 Active Ingredients and Mock Brand Information for the Test

Active Ingredient	Drug Category	Mock brand
Ibuprofen (IBU)	Pain reliever	Hexidvil
Dextromethorphan (DEX)	Cough and cold	Circussin
Ranitidine (RAN)	Anti-diarrhea	Recantac

Each brand had two label designs: (1) Label with FOP (Front of Pack label, FOP) and (2) no FOP (Standard). Label design was crossed with highlighting of critical information at two levels (present and absent). Label design and highlight were crossed for a total of four treatments (FOP with highlight; FOP without highlight; standard with highlight and standard without highlight). This made for a total of twelve unique stimulus where changes could be implemented. Figure 2.1 shows a group of four Circussin treatments as an example. For other two active ingredients, please see Appendix A. In this study, each participant was asked to complete a total of 168 trials- 56 trials for each of the three mock brands; trial changes differed in the information and locations with details following.

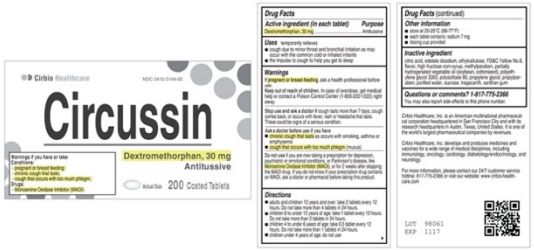


(a) Standard, no highlight



(b) Standard, highlight

(c) Front-of-Pack, no highlight



(d) Front-of-Pack, highlight

Figure 2.1 A Group of Circussin treatments as an example. (a) Standard, no Highlight; (b) Standard, Highlight; (c) Front-of-Pack, no Highlight; (d) Front-of-Pack, Highlight

Three pieces of information were considered as critical to the safe and effective use of OTCs: the active ingredient (AI) which appears on the PDP and in the DFL, drug-drug interactions warnings (DD1, e.g.do not take this drug if you are currently taking aspirin or other blood thinning products which appeared in the DFL and in the PDP only in treatments which included FOPs) and drug-diagnosis interactions warnings (DD2, e.g. do not take this drug if you have been diagnosed with high blood pressure, etc. which appears in the DFL and only in the PDP for treatments which include an FOP). Trials for which changes happened in critical information (AI, DD1 or DD2) were considered as "critical trials" and included in the final analysis.

For both FOP and Standard label formats of the same mock brand, changes to critical information in the DFLs location were identical. Also, the information that was truncated to fit within the confines of the FOP label were changed with the same content and formatting that appeared in the DFL. In other words, changes to critical information in the PDP with the that occurred within the FOP were identical to changes in critical information which occurred in the DFL.

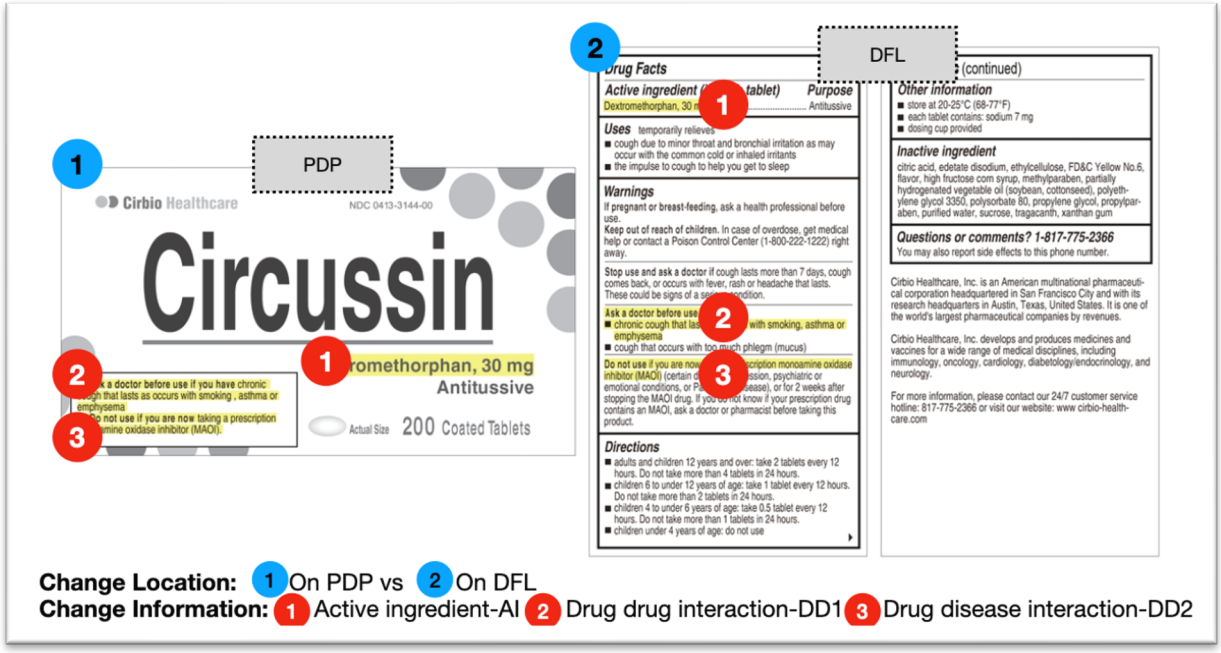


Figure 2.2 An Example of Change Location and Change Information under Highlighting condition of Circussin

It is necessary to note, however, changes to the critical information, active ingredient (AI), had a location confound. Specifically, because the information normally appears on the PDP under commercial conditions it was not included in the novel FOP treatments; additionally, the size of active ingredient information is generally larger than what is presented in the DFL. We honored this size difference to mimic realistic practice. As such, there was a size confound by placement, whereby the presentation of the active ingredient information is larger within the PDP than it is in the DFL. Further, because commercial treatments (represented by our standard

label) do not generally incorporate DD1 or DD2 on the PDP, there was an imbalance of the number of trials by location, whereby, DD1 and DD2 were not tested in the PDP for any standard treatment (either highlighted or non-highlighted conditions). (Figure 2.2 and Figure 2.3)

To prevent participants from preferentially attending the critical information of interest, changes to "non-critical" information were incorporated into the upper half of PDP and the second panel of the DFL to prevent participants from only focusing on specific areas of the package. Therefore, we also designed a total of 36 "non-critical" trials (for both FOP (18 trials) and standard treatments (18 trials)). In sum, each participant was asked to finish 168 trials (56 trials per mock brand) in this change detection test. (Figure 2.3) The trials were randomized to minimize the run order effect.

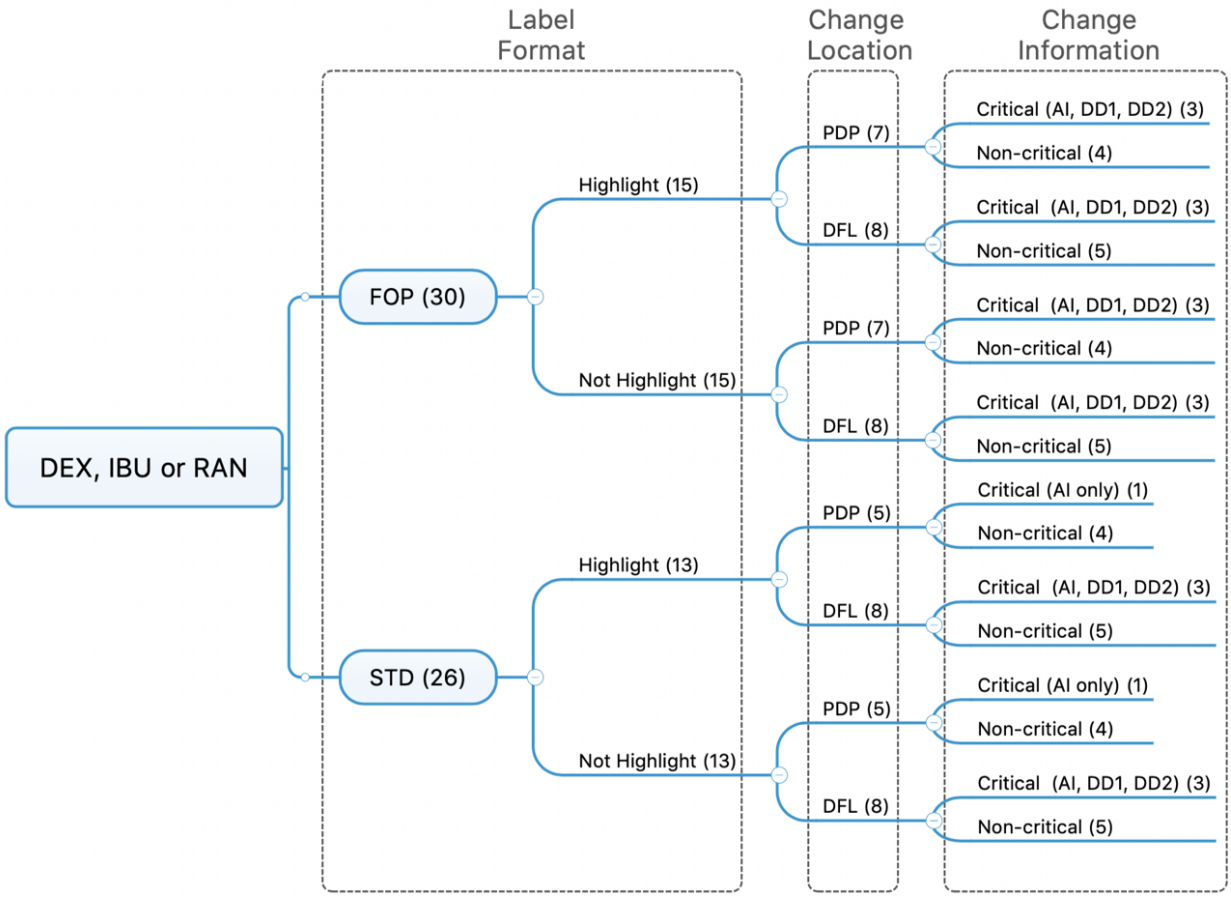


Figure 2.3 Dendrogram of Trial Design for Each Participant

2.3 Participants and Recruitment

In total, 92 participants were recruited via the SONA recruiting system administered by the MSU College of Communication Art and Sciences, as well as circulation of IRB approved flyers (APPENDIX B) and word of mouth advertisement. Participants recruited were eligible if they met the following criteria:

- Be at least 18 years old
- Not be legally blind
- Have used OTC drugs during the past 6 months
- Have NO history of seizure
- Are willing to come the Healthcare, Universal Design, Biomechanics lab (HUB) at the Michigan State University, where the research was conducted.

2.4 Procedures

Upon arrival at the testing lab, participants were provided with a 2-hour parking pass (where applicable) and provided with an IRB-approved consent form to review and sign (APPENDIX C) Participants were informed of their right to stop and opt out any time during the test, and still receive the $25 incentive provided in exchange for participation. Data was recorded by participant number, with no link to participants' identity, and it was protected in a secured storage only accessible by the research team and members of the HRPP upon request.

2.4.1 Demographic survey and Pre-tests

Upon obtaining informed consent, participants were asked to provide basic demographic information, including gender, age, educational level, ethnicity. (APPENDIX D.A) Participants were further characterized using a series of pretests, the details of which are provided in APPENDIX D.

The pre-tests included three standard tests:

(1) *Near-point visual acuity* was characterized using a Bernell vision card (Mishawaka, IN a division of Vision Training Products, INC). (APPENDIX D.B.I) Researchers asked

participants to hold the vision card 16 inches away from their eyes under standard room illumination. Participants were asked to read the lowest (smallest) line on the card they were able to without excessively straining. The lowest line they read completely correctly was scored in accordance with test directives and ranged from 20/20 (lowest line) to 20/800.

(2) The *Rapid Estimate of Adult Literacy in Medicine - Revised (REALM-R)* estimated participant's health literacy. (APPENDIX D.B.II) During the test, participants were asked to read 11 words aloud from the testing card. The first three words are designed as practice trials, and in accordance with the standard, are not scored. The number of words participants read correctly thereafter were scored with a range of 0 to 8. In accordance with standard procedure, scores that are less than 6 are at risk for low health literacy.

(3) Participants' *ability to perceive and differentiate color* (APPENDIX D.C) was also tested with the Tests for Color-blindness (Ishihara, S., 1918), where participants were asked to view 24 color plates and to indicate the number that appeared within each plate, or that no number was present. The ability to decipher the number in the corresponding color plates indicates their ability in color vision whether they are "normal" or at risk for color blindness of "red-green deficiencies" or "total color blindness".

2.4.2 Change Detection Test

After the pretests, participants were assigned to one of the Dell laptop workstations which ran the change detection program with E-prime 3.0 software. Each laptop had an Intel Core i5-7440HQ CPU, 16GB RAM, 238GB memory, a 13-inch 1920*1080 display and 64 bit Windows 10 operating system. Participants were asked to sit in front of the laptop to get ready, while researchers entered the data collected on paper during the previous steps. Keyed information included: participant number, computer number, subject's sex, age, ethnicity, educational level,

native language, near point visual acuity score, REALM-R health literacy score, and color differentiation ability. Once ready, participants were asked to start with the change detection program.

The seizure screening criteria information below was displayed at the beginning of the program to reaffirm participants' eligibility regarding no seizure history.

> *"Do you have a history of seizures? If so, we ask that you do not participate in the experiment. Instead, please inform the experimenter that you are not able to continue. If you have no history of seizures, please hit a button to continue."*

Once participants confirmed they were eligible for the study (affirming no history of seizure), a brief introduction of the experiment appeared along with a welcome for the participant.

> *"Welcome to the experiment! You will see two images separated by a brief blank. The images are identical except for one change. Your task is to detect the change. As soon as you see the change, press the space bar. Then the cursor will appear. Use the mouse to click on the location where the change occurred. The task is timed until you hit the space bar. Using the mouse to indicate the change location is not timed. If you fail to find the change within 18 seconds, the trial will time out. Please hit the space bar to begin a few practice trials..."*

After the introduction page, participants were directed to finish four sample change detection trials to warm up and get familiar with testing operations. Researchers were present to field questions as well as assist with operation and interaction with the program. The warm-up trials were introduced with the following instruction:

> *"Try to find the change that appears in the image, or 'flickering' as quickly as possible. You can indicate that you have found this by hitting the space bar. After this point, you will need to use the mouse to click the area where the change appeared."*

Upon completion of warm up trials (a total of four), participants were asked if they had any questions that the research team could clarify; after which point, formal trials began. Participants could start when they were ready. A total of 168 trials were randomly and evenly

divided into three blocks (56 trials each) (Figure 2.4), each of the blocks had the identical number of trials from a given mock brand. After each block of 56 trials, participants could choose to take a break or continue.

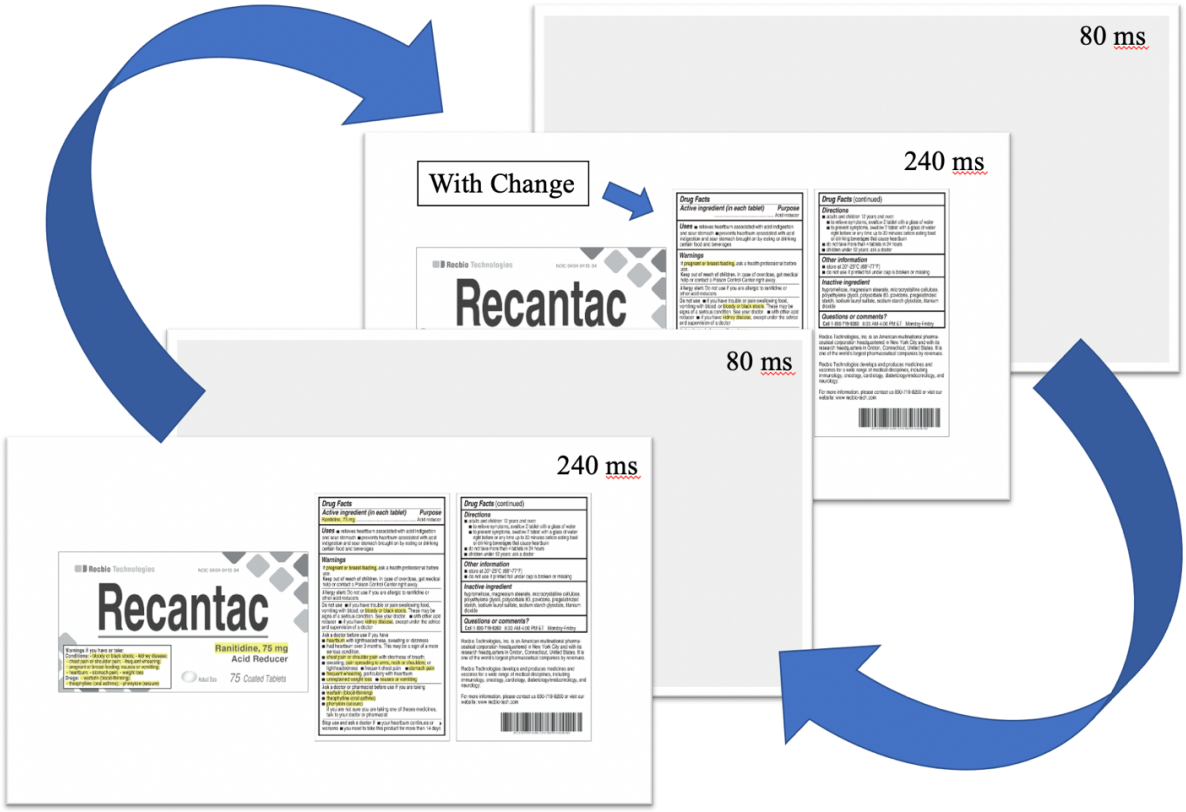


Figure 2.4 Change Detection Image Flickering Cycle for a Trial

2.4.3 Variables and Measurements

To test our hypothesis that an FOP labeling strategy employing highlights is] effective] for garnering attention, we focused on two dependent variables: (1) whether participants that were able to successfully detect changes in critical trials (a binary variable, yes or no) prior to timing out at 18 seconds, and (2) the time to correctly identify the change (a continuous variable, time in seconds) prior to timing out.

The predictor variables included in the final model were label design (Front-of-pack labels or FOP vs standard labels), highlight (content highlighted vs. not highlighted), change

location (changes blinking on the principal display panel PDP vs. changes blinking on the drug facts label DFL), change information (the content information to change or blink: active ingredient, DD1 drug-drug interaction information, or DD2 drug-diagnosis information), ingredients (ranitidine-RAN, ibuprofen-IBU or dextromethorphan-DEX), and covariates involving age, education (below Bachelor, Bachelor or above Bachelor), ethnicity (white or other), sex (male or female)

2.5 Descriptive Statistics

A total of 92 participants (27 male and 65 female) were recruited via the SONA system, all of whom completed our change detection testing. Most of the participants were under 30 years old (60/92, 65.2%), white (62/92, 67.4%) and spoke English as their native language (78/92, 84.8%). The participants were highly educated, with more than two thirds (62/92, 67.4%) of the population reporting receipt of bachelor's degree or higher; this was also supported by REALM-R scores, 87 of the test population (87/92, 94.36%) scored at levels that were not at risk for poor health literacy (REALM-R score ≥ 7). All the testing participants were in the normal range of visual acuity with the corrected near-point visual acuity test score better than 20/40. (92/92, 100%). Regarding the ability to view color, 89 out of 92 participants were in the typical levels with 3 indicated to be at risk for color vision. More details about the demographic information are shown in the following Table 2.2 and Figure 2.5.

Table 2.2 Descriptive Statistics of Change Detection

Characteristic	Value	Number	% Of Total (92)
Sex	Male	27	29.3%
	Female	65	70.7%
Color Differentiation	Deficit	89	96.7%
	Normal	3	3.3%
Education	Below Bachelor	30	32.6%
	Bachelor's degree	36	39.1%
	Above Bachelor	26	28.3%
Ethnicity	Other	30	32.6%
	White	62	67.4%
Health Literacy	5 or lower	5	5.4%
	6 or higher	87	94.6%
Native Language	Other	14	15.2%
	English	78	84.8%
Visual Acuity	20/40 or better	92	100.0%
	Poorer than 20/40	0	0.0%
Age	Mean (Min, Max)	31.55	Min = 19, Max = 67
	Std. Deviation	12.717	

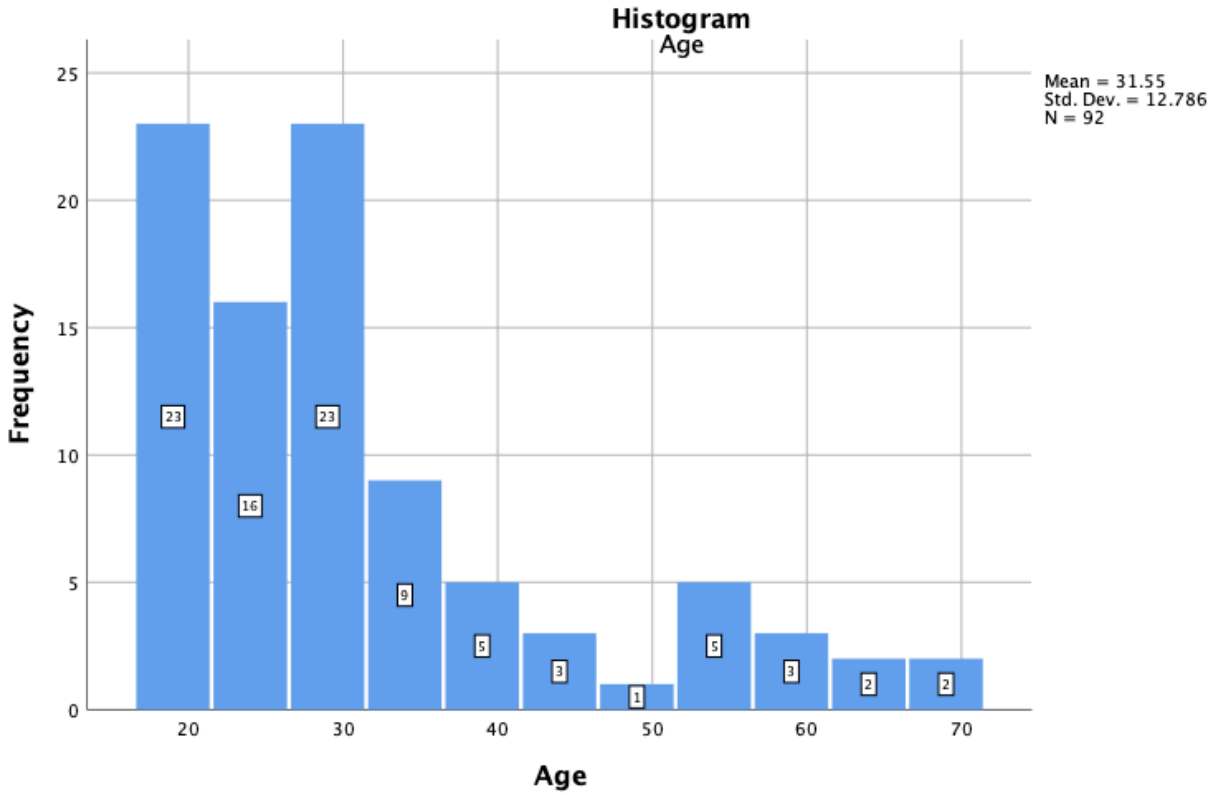


Figure 2.5 Histogram of Age in the Change Detection Test

2.6 Data Analysis and Results in response to Research Question 1

To answer the Research Question 1 (Chapter 2.1), two sets of analyses were conducted. Noticing that the changes which were outside of the FOP were consist of the changes on the DFLs and the changes of AI on the PDPs, we conducted two groups of analyses based on each of the two types of changes respectively. Analyses 1 compared the differences in participants' attention garnering (response accuracy and response time to correct answers) between label layouts (with FOP strategy and the standard layout) when changes occurred in the DFL location. (Figure 2.6.a) And Analyses 2 compared the same differences when the changes were active ingredients on the PDPs. (Figure 2.6.b)

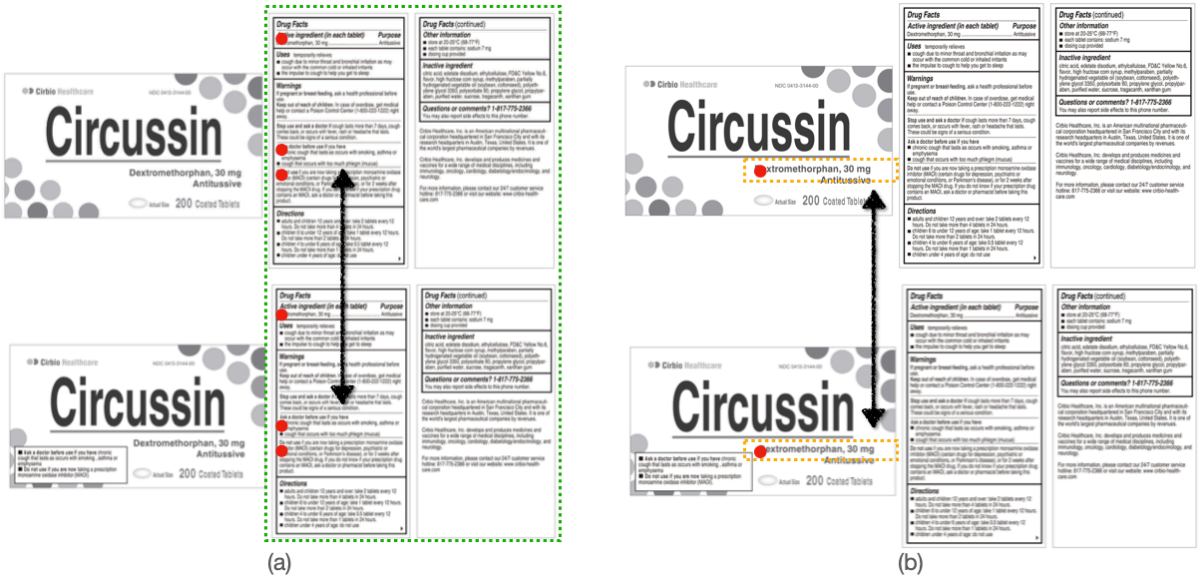


Figure 2.6 Illustration of Two Groups of Analyses in Response to Research Question 1

2.6.1 Analyses 1: When Changes on the DFLs - Change Detection Accuracy

A generalized linear mixed model was fitted to the probability of change detection (in percentage). Only critical trials were analyzed, that is the trials (see Figure 2.3) with changes involving the active ingredient (AI), drug-drug interaction (DD1) or drug-diagnosis (DD2). The binary data of successful change detection prior to time out were transformed and interpreted in

terms of probability (p) of correctly answering trial questions with logit transformation, *ln(p/1-p)*. Tukey's method was used for minoring non-constant variance, and Satterthwaite's method was utilized for adjusting degree of freedom.

As forementioned in Chapter 2.4.3, the predictor variables included in the final model were label design, highlight, change location, change information, ingredients, and covariates involving age, education, ethnicity, sex. All possible 2-way and 3-way interactions across Label type, highlight and change location were also included. All estimated means were back transformed as the original scale of the dependent variable: the probability of successfully detecting changes in percentage.

A total of 5,520 trials/observations (92 subjects x 60 critical trials) were analyzed in this change detection test. Participants successfully detected changed in 4,649 trials (84.2%) and failed to successfully detect prior to timing out for a total of 871 trials (15.8%).

A summation of the results from the statistical analysis is presented in Table 2.3. For the variables of interest, there were significant fixed effects of highlight ($p<0.001$) on response accuracy. Main effects yielded, no evidence of significant effects of Label type ($p=0.179$) or change location ($p=0.139$) on the response accuracy. Two-way interactions were found statistically significant when Label type and change location were crossed ($p=0.023$), and when Label type and highlight were crossed ($p=0.023$). However, due to a 3-way interaction of Label type x highlighting x change location also indicated to statistically significantly impact accuracy, we focus our analysis there. (Table 2.3) The results of other analyses in this model (the effects of top significant effects, and the significant 2-way interactions) beyond the 3-way interactions were archived in Appendix K for readers' interests.

Table 2.3 Fixed Effects of Variables on Change Detection Accuracy

Source	*F*	*df1*	*df2*	*Sig.*
Corrected Model	25.576	16	255	0.000
FOPtype	1.809	1	5503	0.179
Highlight	**169.606**	**1**	**5503**	**0.000**
ChangeLocation	2.188	1	5503	0.139
ChangeInformation	**61.081**	**2**	**5503**	**0.000**
Ingredients	**38.994**	**2**	**5503**	**0.000**
Age	**34.209**	**1**	**73**	**0.000**
Education	2.942	2	83	0.058
Ethnicity	2.163	1	84	0.145
Sex	0.171	1	83	0.681
FOPtype * Highlight	**5.183**	**1**	**5503**	**0.023**
FOPtype * ChangeLocation	**5.152**	**1**	**5503**	**0.023**
*Highlight * ChangeLocation*	0.018	1	5503	0.895
FOPtype * Highlight * ChangeLocation	**4.019**	**1**	**5503**	**0.045**

Probability distribution: Binomial; Link function: Logit; Target: Response

2.6.1.1 Significant 3-way Interaction: Label type x Highlight x Change Location

To interpret this 3-way interaction among Label type x Highlight x Change Location, pairwise comparisons were analyzed, and results are presented in Figure 2.7 and Table 2.4-2.7. As aforementioned in Figure 2.6 (a), the analysis was focused on the comparisons among the trials with changes occurring on the DFLs (as illustrated in green and orange columns respectively in Figure 2.8).

For the effects of highlight, evidence suggested that participants were more likely to detect changes when information was highlighted than those were not. (See Table 2.6). Specifically, when the FOP labeling strategies were applied, highlighting information content increased the probability of detecting changes when the changes were on DFL locations (contrast estimate=0.153, SE=0.02, p=1.89E-14<0.001). Similarly, when the standard labels were presented, highlighting content also increased the probability of detecting changes when the changes were on the DFLs (contrast estimate=0.148, SE=0.019, p=2.73E-14<0.001).

For the effects of FOP types, (Table 2.5) however, no evidence was found that participants performed differently in the trials with FOP labels than the standard ones (contrast estimate=-0.008, SE=0.011, p=0.462) when the content was highlighted. Similarly, no evidence of a statistically significant difference when detecting changes with the trials of FOP labels (ME=0.779, SE=0.022) compared the standard ones (ME=0.792, SE=0.021) (contrast estimate=-0.013, SE=0.021, p=0.533) without highlighting.

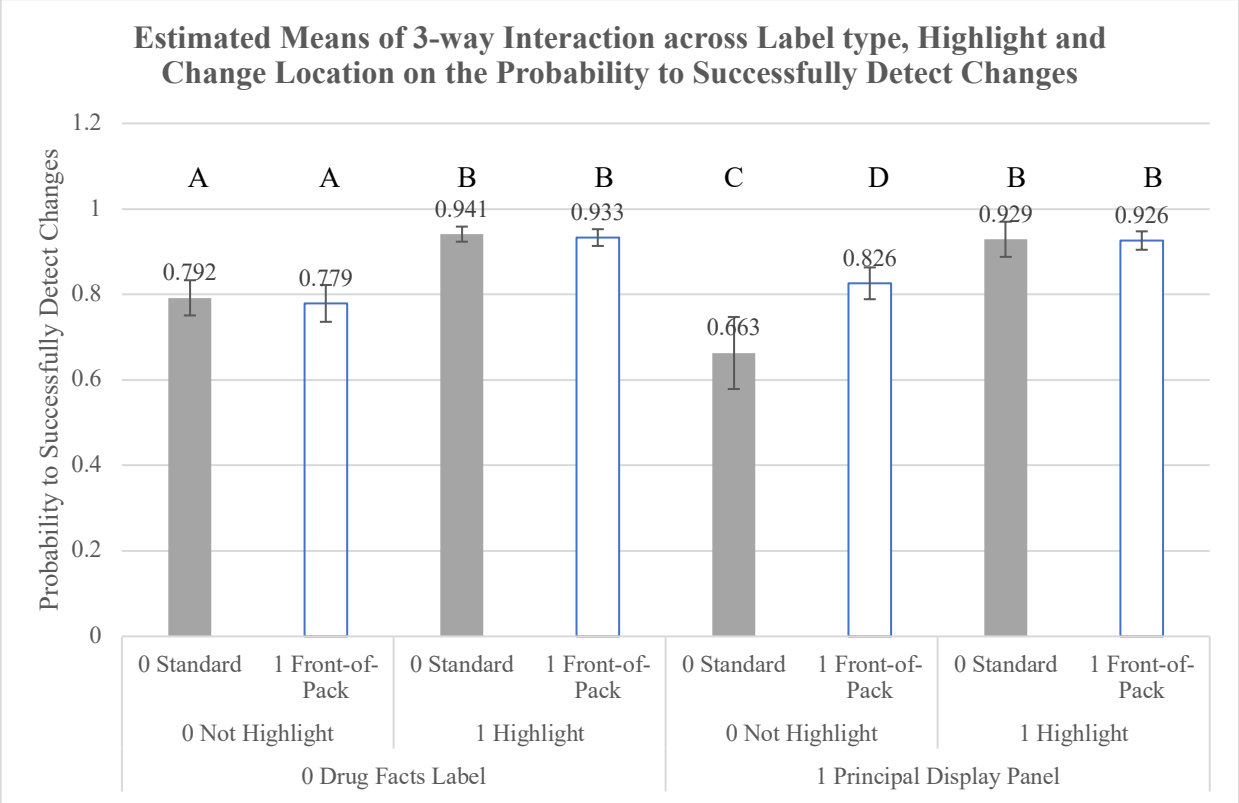


Figure 2.7 Estimated Means of 3-way Interaction across Label type, Highlight and Change Location on the Probability to Successfully Detect Changes

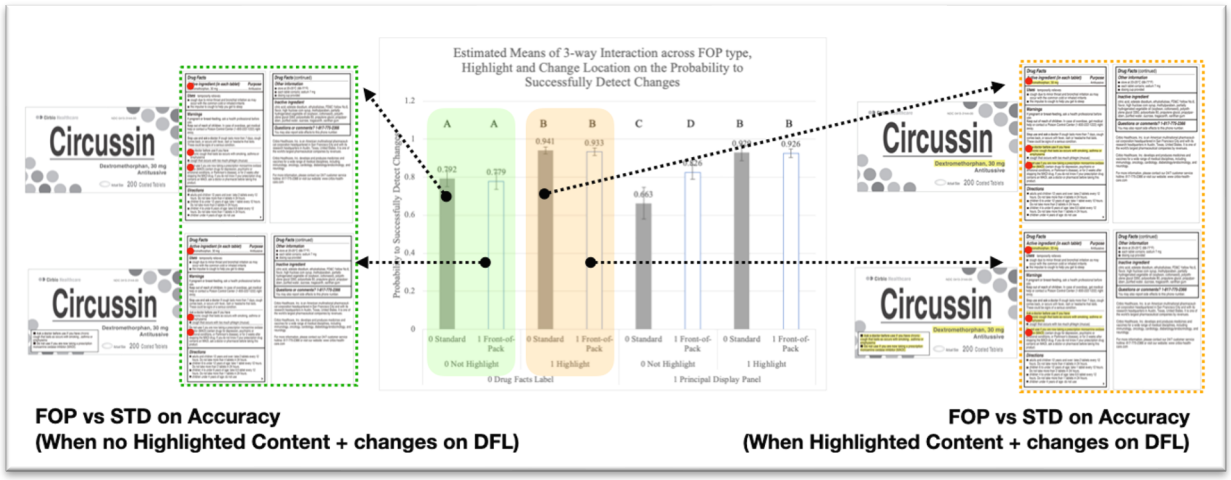


Figure 2.8 Illustration of Comparisons of Accuracy for the Changes on DFLs between different Label types

Table 2.4 Results of Estimated Means of 3-way Interaction across Label type, Highlight and Change Location on the Probability to Successfully Detect Changes

Label type	*Highlight*	*Change Location*	*Mean*	*Std. Error*	*95% Confidence Interval*	
					Lower	Upper
Front-of-Pack	Highlight	Principal Display Panel	0.926	0.011	0.902	0.944
		Drug Facts Label	0.933	0.010	0.91	0.950
	Not Highlight	Principal Display Panel	0.826	0.019	0.785	0.860
		Drug Facts Label	0.779	0.022	0.734	0.819
Standard	Highlight	Principal Display Panel	0.929	0.021	0.875	0.960
		Drug Facts Label	0.941	0.009	0.920	0.956
	Not Highlight	Principal Display Panel	0.663	0.043	0.574	0.742
		Drug Facts Label	0.792	0.021	0.748	0.831

Continuous predictors are fixed at the following values: Age=31.55

Table 2.5 Pairwise Comparisons of Label type on a certain level of Highlight and Change Location

Highlight	*Change Location*	*Label type Pairwise Contrasts*	*Contrast Estimate*	*Std. Error*	*t*	*df*	*Adj. Sig.*	*95% Confidence Interval*	
								Lower	Upper
Highlight	Principal Display Panel	Front-of-Pack vs. Standard	-0.003	0.022	-0.131	5503	8.96E-01	-0.046	0.04
	Drug Facts Label	Front-of-Pack vs. Standard	-0.008	0.011	-0.736	5503	0.462	-0.029	0.013
Not Highlight	Principal Display Panel	Front-of-Pack vs. Standard	0.162	0.042	3.866	5503	0.000	0.08	0.245
	Drug Facts Label	Front-of-Pack vs. Standard	-0.013	0.021	-0.623	5503	5.33E-01	-0.054	0.028

The sequential Bonferroni adjusted significance level is .05. Confidence interval bounds are approximate.

Table 2.6 Pairwise Comparisons of Highlight on a certain level of Label type and Change Location

FOPtype	*Change Location*	*Highlight Pairwise Contrasts*	*Contrast Estimate*	*Std. Error*	*t*	*df*	*Adj. Sig.*	*95% Confidence Interval*	
								Lower	Upper
Front-of-Pack	Principal Display Panel	Highlight vs. Not Highlight	0.1	0.017	5.795	2348	7.74E-09	0.066	0.134
	Drug Facts Label	Highlight vs. Not Highlight	0.153	0.02	7.787	893	1.89E-14	0.115	0.192
Standard	Principal Display Panel	Highlight vs. Not Highlight	0.265	0.044	6.037	5164	1.68E-09	0.179	0.351
	Drug Facts Label	Highlight vs. Not Highlight	0.148	0.019	7.752	808	2.73E-14	0.111	0.186

The sequential Bonferroni adjusted significance level is .05. Confidence interval bounds are approximate.

Table 2.7 Pairwise Comparisons of Change Location on a certain level of Label type and Highlight

Label type	*Highlight*	*Change Location Pairwise Contrasts*	*Contrast Estimate*	*Std. Error*	*t*	*df*	*Adj. Sig.*	*95% Confidence Interval*	
								Lower	Upper
Front-of-Pack	Highlight	Principal Display Panel vs. Drug Facts Label	-0.007	0.012	-0.618	5503	0.537	-0.03	0.016
	Not Highlight	Principal Display Panel vs. Drug Facts Label	0.046	0.02	2.275	5503	0.023	0.006	0.086
Standard	Highlight	Principal Display Panel vs. Drug Facts Label	-0.012	0.022	-0.571	5503	0.568	-0.055	0.03
	Not Highlight	Principal Display Panel vs. Drug Facts Label	-0.129	0.042	-3.065	5503	0.002	-0.212	-0.047

The sequential Bonferroni adjusted significance level is .05. Confidence interval bounds are approximate.

2.6.2 Analyses 1: When Changes on the DFLs - Response Time when successfully detected changes

In addition to the response variable, probability of detecting changes prior to timing out, we also evaluated time to successfully detect the changes as a dependent variable – response time (in seconds for the trials were successful in correctly detecting changes). The data of response time was checked for the validity of normality assumptions prior to statistical analysis. Residual plots and normal probability plots of the original data suggested data transformation was needed. As a result, data were natural log transformed, ln(t), where t represent the original scale of response time. Tukey's method was used for minoring non-constant variance and Satterthwaite's degree of freedom was used to adjust degrees of freedom. A generalized linear mixed model was used to analyze natural log-transformed data.

The predictor variables included in the final model were the same as the previous analysis, namely: Label type, highlight, change location, change information, ingredients, and covariates involving age, education, ethnicity, sex and all possible 2-way and 3-way interactions cross Label type, highlight and change location. All estimated means were back transformed as the original scale of response time in seconds for the visuals that are presented herein.

A summation of the results from the statistical analysis is presented in Table 2.8 For the variables of interest, there were significant fixed effects of Label type ($p<0.001$), highlight ($p<0.001$), change location ($p<0.001$) on response time for correct change detections. Additionally, 2-way interactions were also found statistically significant when Label type and change location were crossed ($p<0.001$), and when Label type and highlight were crossed ($p<0.001$). However, due to a 3-way interaction of Label type x highlight x change location were also found statistically significant on the response time for correct change detections, we focus

our efforts here as effects are mediated by varied inputs. (Table 2.3) The results of other analyses in this model (the effects of top significant effects, and the significant 2-way interactions) beyond the 3-way interactions were archived in Appendix L for readers' interests.

Table 2.8 Fixed Effects of Variables on Response Time for Correct Change Detections

Source	*F*	*df1*	*df2*	*Sig.*
Corrected Model	**336.65**	**16**	**4632**	**0.000**
FOPtype	**23.775**	**1**	**4632**	**0.000**
Highlight	**447.251**	**1**	**4632**	**0.000**
ChangeLocation	**2654.154**	**1**	**4632**	**0.000**
ChangeInformation	**434.532**	**2**	**4632**	**0.000**
Ingredients	**22.823**	**2**	**4632**	**0.000**
Age	3.74	1	4632	0.053
Education	0.147	2	4632	0.863
Ethnicity	0.226	1	4632	0.634
Sex	0.702	1	4632	0.402
FOPtype * Highlight	**7.422**	**1**	**4632**	**0.006**
*FOPtype * ChangeLocation*	0.108	1	4632	0.743
Highlight * ChangeLocation	**29.423**	**1**	**4632**	**0.000**
FOPtype * Highlight * ChangeLocation	**19.697**	**1**	**4632**	**0.000**

Link function: Log Target: Response Time

2.6.2.1 Significant 3-way Interactions: Label type x Highlight x Change Location

To interpret the three-way interaction of label type x Highlight x Change Location was considered, pairwise comparisons were analyzed with results presented in Figure 2.9 and Table 2.9-2.12.

For the effects of highlight, evidence supported the idea that participants spent less time (in average) to correctly detect the changes when content was highlighted shown in the Table 2.10. Specifically, in the presence of front-of-pack labeling strategies, highlighting content shortened the response time to correctly detect changes compared with FOP label trials without highlighted, this was consistent regardless of change location; PDP (contrast estimate=-0.841,

SE=0.084, p<0.001) or DFL (contrast estimate=-1.588, SE=0.175, p=<0.001). Similarly, when the standard labels were presented, highlighting content also shortened the response time to correctly detect changes compared with the trials without highlighted trails, regardless of change location (PDP (contrast estimate=-1.375, SE=0.142, p=<0.001) or DFL) (contrast estimate=-1.183, SE=0.154, p=1.93E-14<0.001).

Table 2.9 Results of Estimated Means of 3-way Interaction across Label types, Highlight and Change Location on Response Time on Correct Change Detection (Seconds)

	Highlight	*Change Location*	*Mean*	*Std. Err.*	*95% Confidence Interval*	
					Lower	Upper
Front-of-Pack	Highlight	Principal Display Panel	2.738	0.148	2.463	3.045
		Drug Facts Label	5.845	0.316	5.257	6.498
	Not Highlight	Principal Display Panel	3.579	0.195	3.217	3.981
		Drug Facts Label	7.433	0.405	6.679	8.272
Standard	Highlight	Principal Display Panel	2.323	0.136	2.071	2.607
		Drug Facts Label	5.55	0.3	4.993	6.17
	Not Highlight	Principal Display Panel	3.698	0.221	3.289	4.159
		Drug Facts Label	6.733	0.367	6.051	7.492

Continuous predictors are fixed at the following values: Age=30.61

For the effects of label types, evidence was found that participants spent longer time to detect changes on the trials with front-of-pack labels compared to trials with standard labels under the following three conditions of highlight and change location (Figure 2.9 and Table 2.11): (1) highlighted and changes occurred to the PDP (contrast estimate=0.415, SE=0.079, p<0.001); (2) highlighted and changes took place on the DFL (contrast estimate=0.294, SE=0.124, p=0.018); (3) no highlights and changes occurred on the DFL (contrast estimate=-3.035, SE=0.222, p=<0.001). However, there was one exception, which suggested no difference in response time between FOP labels and standard ones when the content was highlighted, and changes were on the PDP. (p=0.342>0.05) Under this condition, it is worth noting that the data

of FOP format involved two more change information data compared to that of standard format due to the unbalanced experimental design caused by its own characteristics of FOP labels. (Figure 2.3)

For the effects of change location (Table 2.12), participants generally took less time to identify changes correctly when the changes appeared on the PDPs than those presented on DFLs, no matter the labels were front-of-pack with highlights (contrast estimate=-3.106, SE=0.189, $p<0.001$); or standard with highlights (contrast estimate=-3.227, SE=0.198, $p<0.001$); front-of-pack labels with no highlights (contrast estimate=-3.854, SE=0.243, $p<0.001$); or standard with no highlights (contrast estimate=-3.035, SE=0.222, p=<0.001).

When the changes were on the DFLs as shown in the left half of Figure 2.10, evidence was found that participants spent longer time to detect changes with the FOP label format compared to standard label format. Specifically, when the label content was highlighted, the contrast estimate of response time between FOP layouts and standard layouts was 0.294. (SE=0.124, p=0.018). When not highlighted, the contrast estimate of response time was -3.035. (SE=0.222, p=<0.001)

However, for the changes were on the PDPs as shown in the right half of Figure 2.10, due to the unbalanced experimental design, additional analyses were needed to single out the effects of each change on PDPs. For this reason, we conducted Analyses 2 in the following sections and investigated RQ2 in Chapter 2.7.

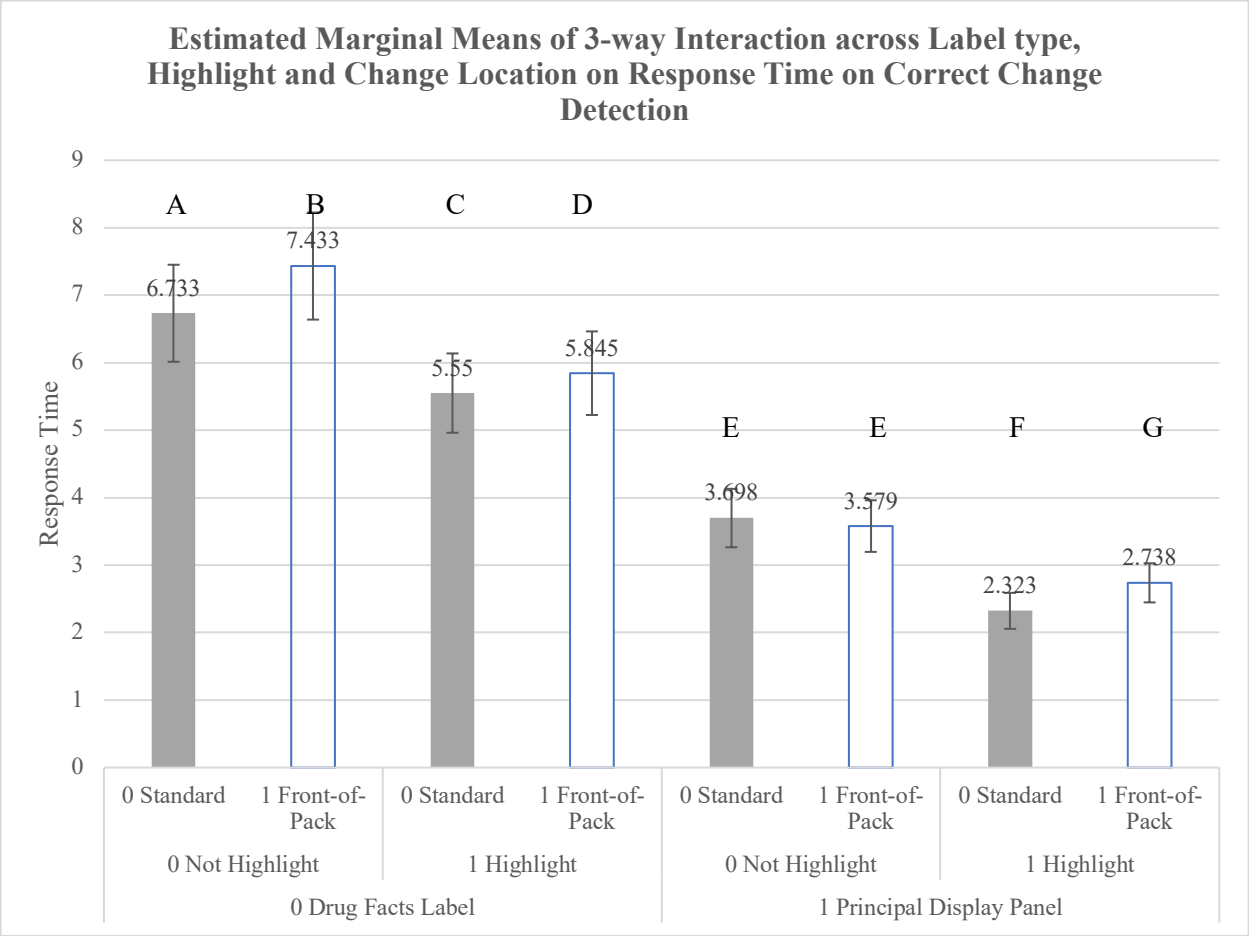


Figure 2.9 Estimated Means of 3-way Interaction across Label type, Highlight and Change Location on Response Time on Correct Change Detection (Seconds)

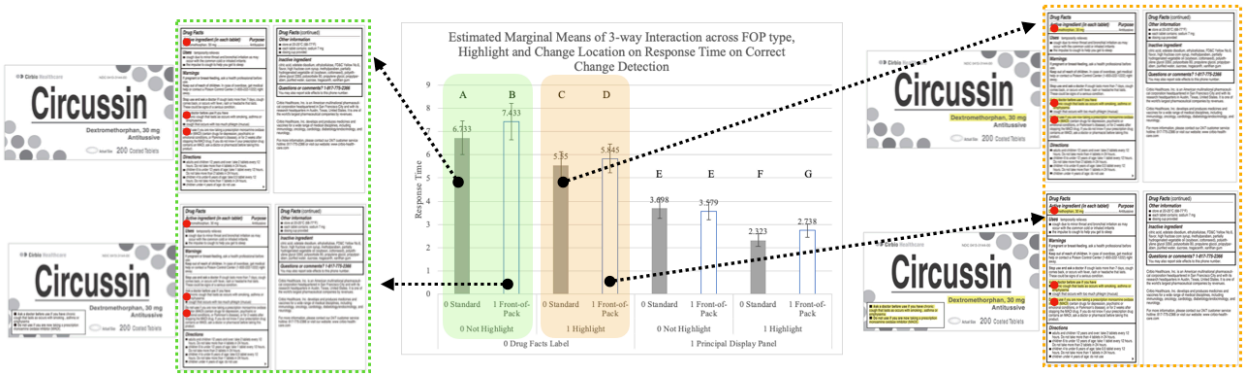


Figure 2.10 Illustration of Comparisons of Response Time to Correct Change Detection for the Changes on DFLs between different Label types.

Table 2.10 Pairwise Comparisons of Highlight on a certain level of Label type and Change Location

Label type	*Change Location*	*Highlight Pairwise Contrasts*	*Contrast Estimate*	*Std. Error*	*t*	*df*	*Adj. Sig.*	*95% Confidence Interval*	
								Lower	Upper
Front-of-Pack	Principal Display Panel	Highlight vs. Not Highlight	-0.841	0.084	-9.952	4632	0.000	-1.006	-0.675
	Drug Facts Label	Highlight vs. Not Highlight	-1.588	0.175	-9.069	4632	0.000	-1.932	-1.245
Standard	Principal Display Panel	Highlight vs. Not Highlight	-1.375	0.142	-9.711	4632	0.000	-1.652	-1.097
	Drug Facts Label	Highlight vs. Not Highlight	-1.183	0.154	-7.68	4632	1.93E-14	-1.485	-0.881

The sequential Bonferroni adjusted significance level is .05. Confidence interval bounds are approximate.

Table 2.11 Pairwise Comparisons of Label type on a certain level of Highlight and Change Location

Highlight	*Change Location*	*Label type Pairwise Contrasts*	*Contrast Estimate*	*Std. Error*	*t*	*df*	*Adj. Sig.*	*95% Confidence Interval*	
								Lower	Upper
Highlight	Principal Display Panel	Front-of-Pack vs. Standard	0.415	0.079	5.228	4632	1.78E-07	0.259	0.571
	Drug Facts Label	Front-of-Pack vs. Standard	0.294	0.124	2.368	4632	0.018	0.051	0.538
Not Highlight	Principal Display Panel	Front-of-Pack vs. Standard	-0.119	0.126	-0.951	4632	0.342	-0.366	0.127
	Drug Facts Label	Front-of-Pack vs. Standard	0.7	0.173	4.055	4632	5.10E-05	0.361	1.038

The sequential Bonferroni adjusted significance level is .05. Confidence interval bounds are approximate.

Table 2.12 Pairwise Comparisons of Change Location on a certain level of Label type and Highlight

Label type	*Highlight*	*Change Location Pairwise Contrasts*	*Contrast Estimate*	*Std. Error*	*t*	*df*	*Adj. Sig.*	*95% Confidence Interval*	
								Lower	Upper
Front-of-Pack	Highlight	Principal Display Panel vs. Drug Facts Label	-3.106	0.189	-16.433	4632	0.000	-3.477	-2.736
	Not Highlight	Principal Display Panel vs. Drug Facts Label	-3.854	0.243	-15.836	4632	0.000	-4.331	-3.377
Standard	Highlight	Principal Display Panel vs. Drug Facts Label	-3.227	0.198	-16.317	4632	0.000	-3.614	-2.839
	Not Highlight	Principal Display Panel vs. Drug Facts Label	-3.035	0.222	-13.645	4632	0.000	-3.471	-2.599

The sequential Bonferroni adjusted significance level is .05. Confidence interval bounds are approximate.

2.6.3 Analyses 2: When Changes were AI on the PDPs - Change Detection Accuracy

A generalized linear mixed model was fitted to the probability of change detection in percentage. Only the critical trials with active ingredient information changed on the PDPs were analyzed (see Figure 2.3). The binary data of successful change detection prior to time out were transformed and interpreted in terms of probability (p) of correctly answering trial questions with logit transformation, *ln(p/1-p)*. Residual method was used to adjust degrees of freedom.

The predictor variables included in the final model were Label type, highlight, ingredients, age, education, ethnicity, language, sex. And 2- interactions between Label type and highlight was also included. All estimated means were back transformed as the original scale of the dependent variable: the probability of successfully detecting changes in percentage.

A total of 1,104 trials/observations (92 subjects x 12 critical trials) were analyzed in this change detection test. Participants successfully detected changed in 1,086 trials (98.3%) and failed in 18 trials (1.6%).

Table 2.13 Fixed Effects of Variables on Change Detection Accuracy

Source	*F*	*df1*	*df2*	*Sig.*
Corrected Model	5.854	11	1092	0.000
Label type	0.232	1	1092	0.630
Highlight	**50.095**	**1**	**1092**	**0.000**
Ingredients	**3.656**	**2**	**1092**	**0.026**
Age	0.195	1	1092	0.659
Education	1.613	2	1092	0.200
Ethnicity	3.18	1	1092	0.075
Language	2.072	1	1092	0.150
Sex	2.782	1	1092	0.096
*Label type * Highlight*	1.332	1	1092	0.249

Probability distribution: Binomial; *Link function: Logit*; *Target: Response*

A summation of the results from the statistical analysis is presented in Table 2.13. For the variables of interest, there were significant fixed effects of highlight ($p<0.001$) on response

accuracy. Specifically, as shown in Table 2.14 and Table 2.15, comparing to the trials with no highlighted content (ME=0.780, SE=0.036), participants detected changes significantly more accurate when the content was highlighted. (ME=0.945, SE=0.013) (contrast estimate=0.166, SE=0.03, p<0.001)

However, no evidence was found the significant differences in response accuracy between different Label types. (p=0.63) Also, no evidence showed the two-way interaction between Label type and highlight. (p<0.001)

Table 2.14 Results of Estimated Means of Highlight on Change Detection Accuracy

Highlight	*Mean*	*Std. Error*	*95% Confidence Interval*	
			Lower	Upper
Highlight	0.945	0.013	0.912	0.967
Not Highlight	0.78	0.036	0.702	0.842

Continuous predictors are fixed at the following values: Age=31.55

Table 2.15 Simple Contrasts of Highlight on Change Detection Accuracy

Highlight Simple Contrasts	*Contrast Estimate*	*Std. Error*	*t*	*df*	*Adj. Sig.*	*95% Confidence Interval*	
						Lower	Upper
Highlight (Not Highlight as base)	0.166	0.03	5.497	1092	4.82E-08	0.107	0.225

The sequential Bonferroni adjusted significance level is .05. Confidence interval bounds are approximate.

2.6.4 Analyses 2: When Changes were AI on the PDPs - Response Time when successfully detected changes

In addition to the response variable of the probability to detect changes, we also evaluated time to successfully detect the changes as a dependent variable – response time (in seconds for the trials that participants correctly detected changes). The data of response time was checked for the validity of normality assumptions and the necessary data transformation was needed. As a result, data were natural log transformed, ln(t), where t represent the original scale

of response time. Residual method was used to adjust degrees of freedom. A generalized linear mixed model was used to analyze natural log-transformed data.

The predictor variables included in the final model were the same as the previous analysis, namely: Label type, highlight, ingredients, age, education, ethnicity, language, sex and the 2-way interaction cross Label type and highlight. All estimated means were back transformed as the original scale of response time in seconds.

Table 2.16 Fixed Effects of Variables on Response Time to Correctly Change Detection

Source	*F*	*df1*	*df2*	*Sig.*
Corrected Model	7.852	11	949	0.000
Label type	**8.023**	**1**	**949**	**0.005**
Highlight	**66.829**	**1**	**949**	**0.000**
Ingredients	2.7	2	949	0.068
Age	0.448	1	949	0.504
Education	0.046	2	949	0.955
Ethnicity	0.108	1	949	0.742
Language	0.786	1	949	0.375
Sex	3.772	1	949	0.052
*Label type * Highlight*	0.403	1	949	0.526

Probability distribution: Gamma; *Link function: Log; Target: Response Time*

A summation of the results from the statistical analysis is presented in Table 2.16. There was significant fixed effect of highlight ($p<0.001$) on response time to correct change detections. Specifically, as shown in Table 2.17 and Table 2.18, comparing to the trials with no highlighted content (ME=4.316, SE=0.272), participants spent significantly less time when the content was highlighted. (ME=2.634, SE=0.159) (contrast estimate=-1.683, SE=0.233, $p<0.001$)

Additionally, the fixed effect of Label type ($p<0.001$) was also found significantly on response time. Specifically, comparing to the trials with standard label layout (ME=3.095, SE=0.191), participants spent significantly longer time when the content was highlighted. (ME=3.673, SE=0.226) (contrast estimate=0.577, SE=0.207, $p=0.005<0.01$)

Table 2.17 Estimated Means of Label type and Highlight on Response Time to Correctly Change Detection

	Mean	*Std. Error*	*95% Confidence Interval*	
			Lower	Upper
Label type				
Front-of-Pack	3.673	0.226	3.254	4.145
Standard	3.095	0.191	2.742	3.494
Highlight				
Highlight	2.634	0.159	2.34	2.964
Not Highlight	4.316	0.272	3.814	4.885

Continuous predictors are fixed at the following values: Age=31.35

Table 2.18 Simple Contrasts of Label type and Highlight on Change Detection Accuracy

Simple Contrasts	*Contrast Estimate*	*Std. Error*	*t*	*df*	*Adj. Sig.*	*95% Confidence Interval*	
						Lower	Upper
Label type	**(Standard as comparing base)**						
Front-of-Pack	0.577	0.207	2.794	949	0.005	0.172	0.983
Highlight	**(Not Highlight as comparing base)**						
Highlight	-1.683	0.233	-7.221	949	1.06E-12	-2.14	-1.225

The sequential Bonferroni adjusted significance level is .05. Confidence interval bounds are approximate.

2.7 Data Analyses and Results in response to Research Question 2

To answer the Research Question 2 (Chapter 2.1), another group of analyses were conducted to compare the differences in participants' attention garnering (response accuracy and response time to correct answers) for the information (DD1 and DD2) changed inside of the FOP box (on PDP) to the same type of information changed outside of the FOP box (on DFL), as shown in Figure 2.11. The following sections detail the results of these analyses.

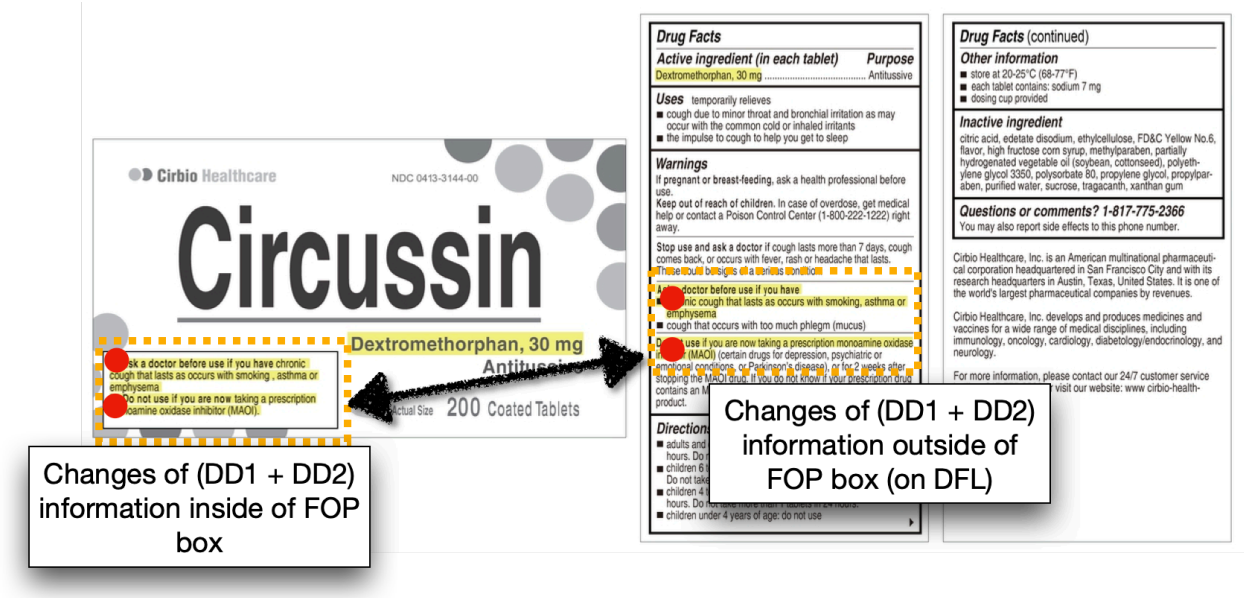


Figure 2.11 Illustration of the Group of Analyses in response to Research Question 2

2.7.1 Change Detection Accuracy

A generalized linear mixed model was fitted to the probability of change detection in percentage. Only critical trials with the changes of DD1 and DD2 were analyzed (see Figure 2.3) The binary data of successful change detection prior to time out were transformed and interpreted in terms of probability (p) of correctly answering trial questions with logit transformation, $ln(p/1-p)$. Residual method was used for adjusting degree of freedom.

As forementioned in Chapter 2.4.3, the predictor variables included in the final model were highlight, change location, ingredients, age, education, ethnicity, language, sex and possible 2-way interaction across highlight and change location. All estimated means were back transformed as the original scale of the dependent variable: the probability of successfully detecting changes in percentage.

A total of 2,208 trials/observations (92 subjects x 24 critical trials) were analyzed in this change detection test. Participants successfully detected changed in 2,068 trials (93.6%) and failed to successfully detect changes prior to timing out in a total of 140 trials (6.3%).

Table 2.19 Fixed Effects of Variables on Change Detection Accuracy

Source	*F*	*df1*	*df2*	*Sig.*
Corrected Model	16.345	11	2196	0.000
Highlight	**83.745**	**1**	**2196**	**0.000**
Change Location	**7.768**	**1**	**2196**	**0.005**
Ingredients	**26.009**	**2**	**2196**	**0.000**
Age	**38.332**	**1**	**2196**	**0.000**
Education	4.032	2	2196	0.018
Ethnicity	1.9	1	2196	0.168
Language	0.724	1	2196	0.395
Sex	0.288	1	2196	0.592
Highlight * Change Location	**7.768**	**1**	**2196**	**0.005**

Probability distribution: Binomial; Link function: Logit; Target: Response

A summation of the results from the statistical analysis is presented in Table 2.19. For the variables of interest, there were significant fixed effects of highlight ($p<0.001$), change location ($p=0.005$), ingredients ($p<0.001$) and age ($p<0.001$) on response accuracy. A two-way interaction was also found statistically significant when highlight and change location were crossed ($p=0.005$).

To interpret this 2-way interaction, pairwise comparisons were conducted, and the results were plotted in Figure 2.12. Under the trials with the critical information was not highlighted, evidence was found that participants detected the changes of information inside of FOP box (DD1 and DD2 information changed on PDP) (ME=0.793, SE=0.028) more accurate than the changes of the same information outside of FOP box (DD1 and DD2 information changed on DFL). (ME=0.657, SE=0.036) (contrast estimate=0.136, SE=0.03, $p<0.001$) However, when highlighted, no evidence for the difference on accuracy between the changes inside of FOP box (ME=0.895, SE=0.018) and the changes outside of FOP box. (ME=0.895, SE=0.018) (contrast estimate = 5.55E-16, $p=1.000$) (Table 2.21 and Figure 2.11)

For the effects of highlighting, evidence showed by highlighting the critical health information increased the detection accuracy than the trials with no highlighting with the contrast estimates of 0.102 (SE=0.024, p<0.001) when the changes of DD1 and DD2 information occurred on the PDPs. Similarly, participants responded more accurately in the trials with highlights than those were not, when the changes of DD1 and DD2 information occurred on the DFLs. (contrast estimates = 0.239, SE=0.031, p<0.001).

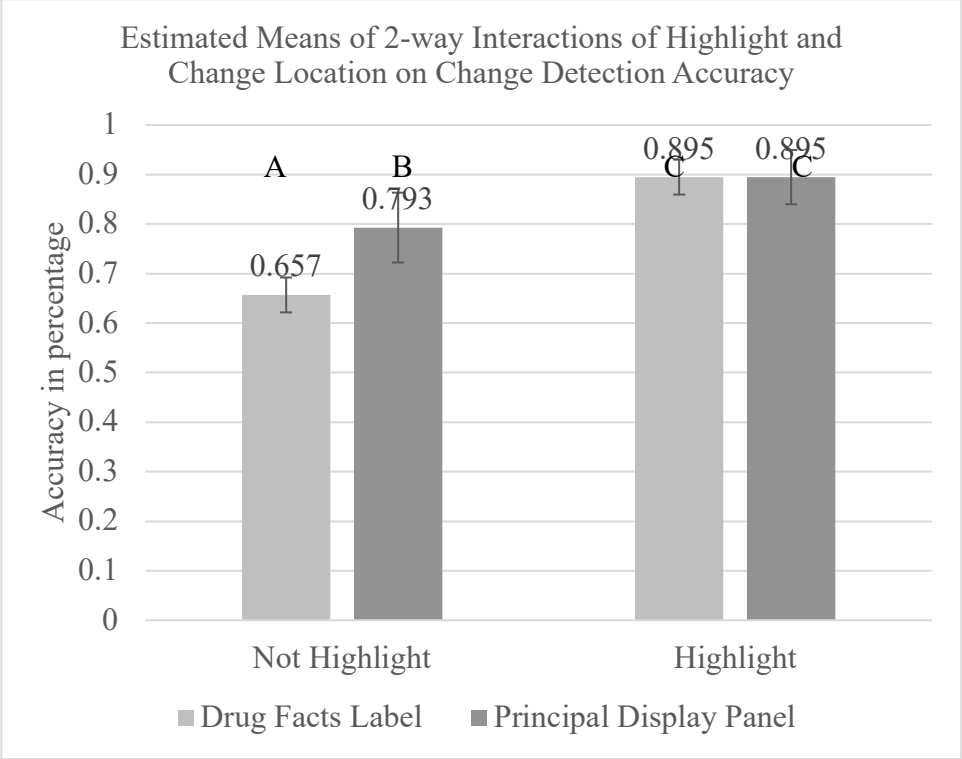


Figure 2.12 Estimated Means of 2-way Interactions of Highlight and Change Location on Response Accuracy Location on Response Accuracy

Table 2.20 Estimated Means of 2-way Interaction between Highlight and Change Location on Change Detection Accuracy

Highlight	*Change Location*	*Mean*	*Std. Error*	*95% Confidence Interval*	
				Lower	Upper
Highlight	Principal Display Panel	0.895	0.018	0.854	0.926
	Drug Facts Label	0.895	0.018	0.854	0.926
Not Highlight	Principal Display Panel	0.793	0.028	0.732	0.842
	Drug Facts Label	0.657	0.036	0.582	0.724

Continuous predictors are fixed at the following values: Age=31.55

Table 2.21 Pairwise Comparisons of 2-way Interaction between Highlight and Change Location on Change Detection Accuracy (Change Location Contrasts)

Highlight	*Change Location Pairwise Contrasts*	*Contrast Estimate*	*Std. Error*	*t*	*df*	*Adj. Sig.*	*95% Confidence Interval*	
	(Drug Facts Label as comparing base)						Lower	Upper
Highlight	Principal Display Panel	5.55E-16	0.019	2.98E-14	2196	1.000	-0.037	0.037
Not Highlight	**Principal Display Panel**	**0.136**	**0.03**	**4.555**	**2196**	**5.53E-06**	**0.078**	**0.195**

The sequential Bonferroni adjusted significance level is .05. Confidence interval bounds are approximate.

Table 2.22 Pairwise Comparisons of 2-way Interaction between Highlight and Change Location on Change Detection Accuracy (Highlight Contrasts)

Change Location	*Highlight Pairwise Contrasts*	*Contrast Estimate*	*Std. Error*	*t*	*df*	*Adj. Sig.*	*95% Confidence Interval*	
	(Not Highlight as comparing base)						Lower	Upper
Principal Display Panel	**Highlight**	**0.102**	**0.024**	**4.198**	**2196**	**2.80E-05**	**0.055**	**0.15**
Drug Facts Label	**Highlight**	**0.239**	**0.031**	**7.718**	**2196**	**1.78E-14**	**0.178**	**0.299**

The sequential Bonferroni adjusted significance level is .05. Confidence interval bounds are approximate.

2.7.2 Response Time when successfully detected changes

In addition to the response variable of the probability to detect changes, we also evaluated time to successfully detect the changes as a dependent variable – response time (in seconds for the trials that participants correctly detected changes). The data of response time was checked for the validity of normality assumptions and the necessary data transformation was needed. As a result, data were natural log transformed, *ln(t)*, where t represent the original scale of response time. Residual method was used to adjust degrees of freedom. A generalized linear mixed model was used to analyze natural log-transformed data.

The predictor variables included in the final model were the same as the previous analysis, namely: highlight, change location, ingredients, age, education, ethnicity, language, sex

and the 2-way interaction cross highlight and change location. All estimated means were back transformed as the original scale of response time in seconds.

Table 2.23 Fixed Effects of Variables on Response Time to Correctly Change Detection

Source	*F*	*df1*	*df2*	*Sig.*
Corrected Model	45.529	11	1768	**0.000**
Highlight	**54.793**	**1**	**1768**	**0.000**
Change Location	**404.674**	**1**	**1768**	**0.000**
Ingredients	**13.987**	**2**	**1768**	**0.000**
Age	**18.455**	**1**	**1768**	**0.000**
Education	**6.065**	**2**	**1768**	**0.002**
Ethnicity	0.906	1	1768	0.341
Language	0.867	1	1768	0.352
Sex	**4.258**	**1**	**1768**	**0.039**
*Highlight * Change Location*	0.253	1	1768	0.615

Probability distribution: Gamma; Link function: Log; Target: Response Time

A summation of the results from the statistical analysis is presented in Table 2.23. For the variables of interest, there were significant fixed effects of highlight ($p<0.001$), change location ($p=0.005$), ingredients ($p<0.001$), age ($p<0.001$), education ($p=0.002$) and sex ($p=0.039<0.05$) on response time to correct change detections. Specifically, as shown in Table 2.24 and Table 2.25, comparing to the trials with no highlighted content (ME=7.086, SE=0.258), participants spent significantly less time when the content was highlighted. (ME=5.559, SE=0.193) (contrast estimate=-1.526, SE=0.216, $p<0.001$)

Additionally, the fixed effect of change location ($p<0.001$) was also found significantly on response time. Specifically, under the FOP layout, participants spent less time to correctly detect the DD1 and DD2 changes inside of FOP box (on PDPs) (ME=4.514, SE=0.159) than when change to DD1 and DD2 took place on the DFL. (ME=8.727, SE=0.313) (contrast estimate=-4.21, SE=0.257, $p<0.001$)

Table 2.24 Estimated Means of Highlight and Change Location on Response Time to Correctly Change Detection

	Mean	*Std. Error*	*95% Confidence Interval*	
			Lower	Upper
Highlight				
Highlight	5.559	0.193	5.194	5.95
Not Highlight	7.086	0.258	6.597	7.611
Change Location				
Principal Display Panel	4.514	0.159	4.213	4.837
Drug Facts Label	8.727	0.313	8.134	9.363

Continuous predictors are fixed at the following values: Age=31.35

Table 2.25 Simple Contrasts of Highlight and Change Location on Response Time to Correctly Change Detection

Simple Contrasts	*Contrast Estimate*	*Std. Error*	*t*	*df*	*Adj. Sig.*	*95% Confidence Interval*	
						Lower	Upper
Highlight	(Not Highlight as comparing base)						
Highlight	-1.526	0.216	-7.078	1768	2.10E-12	-1.949	-1.103
Change Location	(Drug Facts Label as comparing base)						
Principal Display Panel	-4.213	0.257	-16.399	1768	0.000	-4.717	-3.709

The sequential Bonferroni adjusted significance level is .05. Confidence interval bounds are approximate.

2.8 Discussion of this Chapter

2.8.1 Conclusions

From the results of Analyses 1 (Chapter 2.6.1-2) and Analyses 2 (Chapter 2.6.3-4), we could draw conclusions in response to the Research Question 1 (Chapter 2.1) as follows: (1) the strategy of highlighting the critical information on OTC packaging increased change detection accuracy and resulted in less response time to detect changes correctly. Both of these supon port the notion that highlighting increased to critical information. (2) Comparisons of performance to changes that occurred outside of the FOP did not yield evidence that an FOP interfered with people's accuracy regarding their attention to other information (Figure 2.6); no evidence suggested differences in response accuracy between the presence of FOP layout and standard

layout. That said, response time was impacted statistically by the presence of an FOP; specifically, participants took more time to detect changes that occurred outside of the FOP when an FOP layout as compared with treatments involving the standard layout.

From the results of the data analyses (Chapter 2.7) in response to the Research Question 2 (Chapter 2.1), we could draw conclusions as follows: (1) as supported by the previous analyses, highlighting positively influenced both accuracy and response time (2) when the DD1 and DD2 information changes (Figure 2.11). Specifically, the information changes inside of the FOP box (on PDP) increased attention (accuracy and was faster in the presence of highlighting) garnering comparing to the same type of information changed outside of the FOP box (on DFL).

2.8.2 Discussion and Implications

In this study, we explored the potential effects of two OTC labeling strategies, namely the FOP labels and highlighting. By testing the attention garnering abilities to critical information via change detection test, the results suggested the promise of highlighting and FOP strategy as means to improve consumers' early stages of information processing on OTC products. It is worth noting that the FOP strategy was effective (as indicated by the enhanced accuracy) and efficient (as indicated by improved time) at garnering attention to to critical information inside of the FOP box on the PDP. In other words, the presence of FOP box could prolong participants' processing time on the information outside of the FOP box. One potential explanation for the longer response time was the distraction effects of FOP box on the information outside of the box. It is reasonable that more attention attracted by the information inside of the FOP box could delay participants' focus on other information outside of the box. Moreover, it could also be possible that the presence of FOP box could "remind" participants to be more cautious and careful to read the drug information on the OTC packaging in general. Regardless which

explanations above were supported, the labeling strategies of highlighting and FOP were effective and efficient to be used to attract consumers' attention.

2.8.3 Limitations and Future work

Despite the meaningful results from this study, it has many limitations regarding generalizability in the research design as well as the FOP labeling strategy itself.

Firstly, this change detection test was limited by generalizability. This study was based on limited drug ingredients and packaging types. Only three drug ingredients and one regular folding carton type in packaging were used to present the broader categories and various types of packaging format in the OTC industries.

Secondly but most importantly, the limitations came from the FOP labeling strategy itself. Since the area of FOP box limited the amount of information to be placed in the box, the research designers in this study had to decide which warning information to be prioritized (DD1 and DD2) and be shown inside of the FOP box based on their own knowledge and understanding. The static information in the FOP box, however, could be critical and helpful for a specific portion of population, but not for all, because different patients had different needs. Therefore, for the future improvement, the personalized labeling strategy in the next chapter could be an answer to this limitation.

Chapter 3 Framework of Personalized Labeling Strategy and Development of its User Interface (Aim 2)

3.1 Background and the Concept of Personalized FOP Labeling

Drug Facts Labels (DFL) were intended to provide a systematized, consumer-focused method for displaying OTC drug information when they became required in the 1990s. However, the DFL has been criticized through the years for its small print (Trivedi et al., 2014), crowded format, and complex wording (Catlin & Brass, 2018). In short, the DFL needs to be optimized to better communicate the information required for safe and effective use of a drug to those engaged in self-care public, particularly those at risk of ADRs.

To do this, several types of information must be processed (through the various stages) by a variety of consumers. Communication goals include: (1) provision of information required to determine if the drug is right for their condition based on health history, e.g., "Ask a doctor before use if: 'The stomach bleeding warning applies to you." (2) provision of information which warns consumers to stop if potential adverse event develops, or if any significant changes in consumers' conditions develop, e.g., "Stop use and ask a doctor if: 'Pain gets worse or lasts more than 10 days'". (3) provision of the information directing consumers on how to use the product correctly, e.g., "Directions: 'Take 1 tablet every 2 to 3 h while symptoms persist'".

All information required by the DFL, including these, are bound to formatting requirements dictated by law. However, even a well-designed DFL is ineffective if the contents are not read by consumers (early-stage processing- attention). Research has suggested this to be an issue. King et al., (2011) suggested that only 48% of subjects stated they always read the usage instructions on OTC pain relievers, and Cryer et al., (2016) indicates that only 42% of subjects stated they read the OTC label entirely during their first time taking a product. The

multiple communication goals of the DFL are bound to formatting requirements dictated by law. However, even a well-designed DFL could be ineffective if the contents were not read by consumers (early stage early-stage processing- attention). For instance, King et al., (2011) suggested that only 48% of subjects stated they always read the usage instructions on OTC pain relievers, and Cryer et al., (2016) indicates that only 42% of subjects stated they read the OTC label entirely in the first time taking a product.

FOP labeling has been proposed and employed as a design strategy intended to assist consumers during product selection for food products. FOPs have been classified as directive, non-directive and semi-directive based on the level of "directiveness" in a system Hodgkins et al., (2012). The "directiveness" is defined as the extent to which the labeling provides the health information associated with a food product. FOP labeling has been proposed as a design strategy intended to assist consumers during product selection. FOPs have been classified as directive, non-directive and semi-directive based on the level of "directiveness" in a system Hodgkins et al., (2012). The "directiveness" is defined as the extent to which the labeling provides the health information associated with a food product.

- "Non-directive" labels list nutrition information and leave the work of healthfulness interpretation to consumers. "Non-directive" labeling examples include: the Guideline Daily Amount (GDA) label in the EU, Facts-Up-Front design, and Nutrition Facts Panel (NFP) in the US; (Figure 3.1.A)
- "Semi-directive" labeling is based on the concept of non-directive labels but overlaid with symbol, icon, color, or other qualitative assessments representing judgment. In the case of food products, semi-directive labels mark the degree "healthfulness" related to

key nutrition components with different levels of color or symbols, e.g., Traffic Light Labeling depicted overlaying the GDA; (Figure 3.1.B)

- "Directive" labels, however, summarize and provide the overall healthfulness of a product directly. For example, the Smart Choices Label in the US (which is now defunct) with green check mark design, Nordic Keyhole the Nordic countries with green keyhole shape. (Figure 3.1.C)

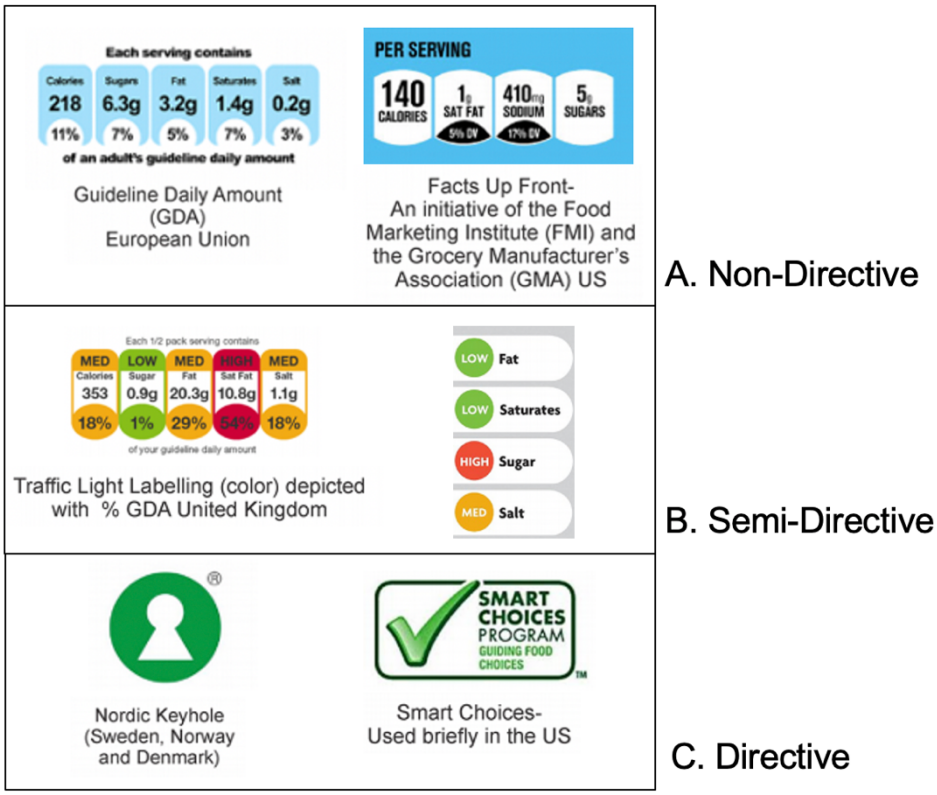


Figure 3.1 Example of Non-directive, Semi-directive, and Directive Front-of-Pack Nutrition Labels

Icons and symbols provide an alternative way to text for communicating information and generally provide an overall summation regarding some aspect of the product. For OTC labeling, a qualitative study using focus groups (King et al., 2011) indicated that consumers supported using icons on OTC packaging to help them to identify the presence of the active ingredient, acetaminophen, which has been identified to have a narrow margin between therapeutic and

problematic dosage. (Figure 3.2) Shiffman et al., (2016) comparative study suggests that participants preferred icons directly connected to an active ingredient as well. (Figure 3.3) Of note, the icons in the previous studies were designed specifically for a given active ingredient, thereby requiring consumers to be educated as to the meaning of a particular icon in advance of its widespread use. As such, we would characterize them as a "semi-directive" label because consumers must utilize their knowledge related to the maximum dosage to interpret the appropriate course of action for their use.

Figure 3.2 Most Preferred Icon for Warning of Acetaminophen (Left) and Message for Maximum Dose (Right). (King et al., 2011)

Figure 3.3 Icons for Acetaminophen as the Active Ingredient. Shiffman et al. (2016)

The aim of the present work is to explore the feasibility of applying a type of "directive" FOP label to OTCs to facilitate their safe and effective use. This "directive" label would employ an FOP label strategy, specifically pointing out criteria relevant for the individual. For an OTC selection by self-medicating consumers, this criterion would be specified as whether this medication is "appropriate to use" (yes/no). This level of customization to labeling requires the labels to be "personalized" based on each individual consumer's health history and current mediations. In other words, the content of the "directive" label is flexible, adapting on a case-by-case basis to meet the goal of personalization. It requires a data-driven system to provide decision support dynamically and response quickly. Traditional printing methods do not enable

this personalized strategy. Hence, it is necessary to leverage digital methods, such as augmented reality technologies, to realize the idea.

3.1.1 Decision Support Systems (DSS)

The framework of a decision support system (DSS) (Sprague, 1980; Pelizaro & Mcdonald, 2006) (Figure 3.4) suggests three main components: (1) the user interface, which helps the end-user interact and communicate with the system; (2) the knowledge base, or database management system, which manages data from both internal sources and external sources, (3) the model management system, which stores algorithms and models used in the decision-making process.

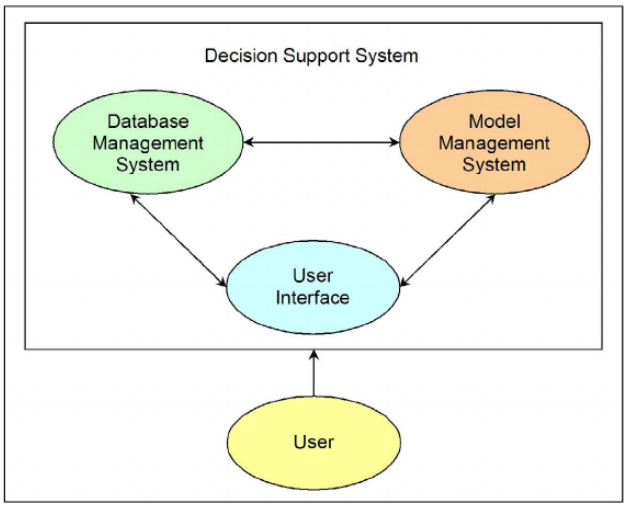


Figure 3.4 A Framework for Decision Support System with Major Components

3.1.2 Clinical Decision Support System (CDSS)

For a decision support system for self-medication, the closest system we can utilize for reference is a clinical decision support system (CDSS). (Sutton et al., 2020) A CDSS is intended to improve healthcare delivery by enhancing medical decisions using targeted clinical knowledge, patient information, and other health information. For the differences in the model management system, CDSS are frequently classified as knowledge-based (expert systems) or

non-knowledge based (algorithm and data-driven systems). (Berner, 2007) In knowledge-based systems, rules are created, with the system retrieving data to evaluate the rule, and return a corresponding action or output. Rules can be made using literature-based, practice-based, or patient-directed evidence. In non-knowledge-based systems, the decision leverages artificial intelligence (AI), machine learning (ML), or statistical pattern recognition from the data source, rather than being programmed or coded to follow expert medical knowledge. (Figure 3.5)

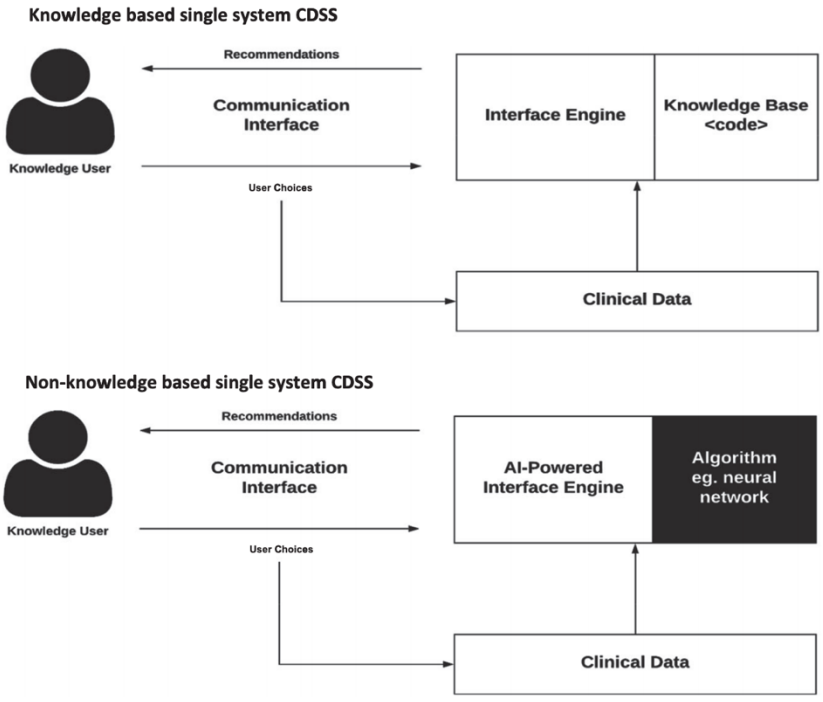


Figure 3.5 Diagram of Key Interactions in Knowledge-based and Non-knowledge based CDSS. (Sutton et al., 2020)

A CDSS can provide a vast range of functions, including diagnostics, alarm systems, disease management, prescription (Rx), drug control. Core and common function of CDSS include reducing medication errors, including common and preventable drug-drug interactions (DDI). One such application is represented by drug safety software; that is, software which includes safeguards for dosing, duplication of therapies, and DDI checking. (Helmons et al.,

2015) A CDSS can provide a vast range of functions, including diagnostics, alarm systems, disease management, prescription (Rx), drug control. One core and common function of CDSS is to reduce medication errors, including common and preventable drug-drug interactions (DDI). One such application is represented by drug safety software; that is, software which includes safeguards for dosing, duplication of therapies, and DDI checking. (Helmons et al., 2015)

3.2 Objectives

The goal of the Aim 2 was to discuss the framework of a personalized FOP labeling concept and to provide a proof of concept regarding the ability to create an AR image related to a personalized FOP labeling strategy.

3.3 Framework of Self-Medication Decision Support System

As suggested previously, the concept of a personalized labeling requires a data-driven system to provide decision support dynamically and quickly. A framework for a self-medication decision support system (SDSS) was presented to enhance understanding.

The framework of an SDSS (depicted in Figure 3.6) included three major components: (1) The user interface -employing augmented reality which overlayed the personalized response in the form of virtual content onto real-world content (the OTC product being considered). (2) The real-world content (the PDP of an OTC medication) was captured by the camera of a device, e.g., iPhone. (3) The virtual content (personalized labeling) was generated according to the decisions of appropriateness for use (shown as a “directive icon”) and was specific to the medication under consideration and the user’s health data.

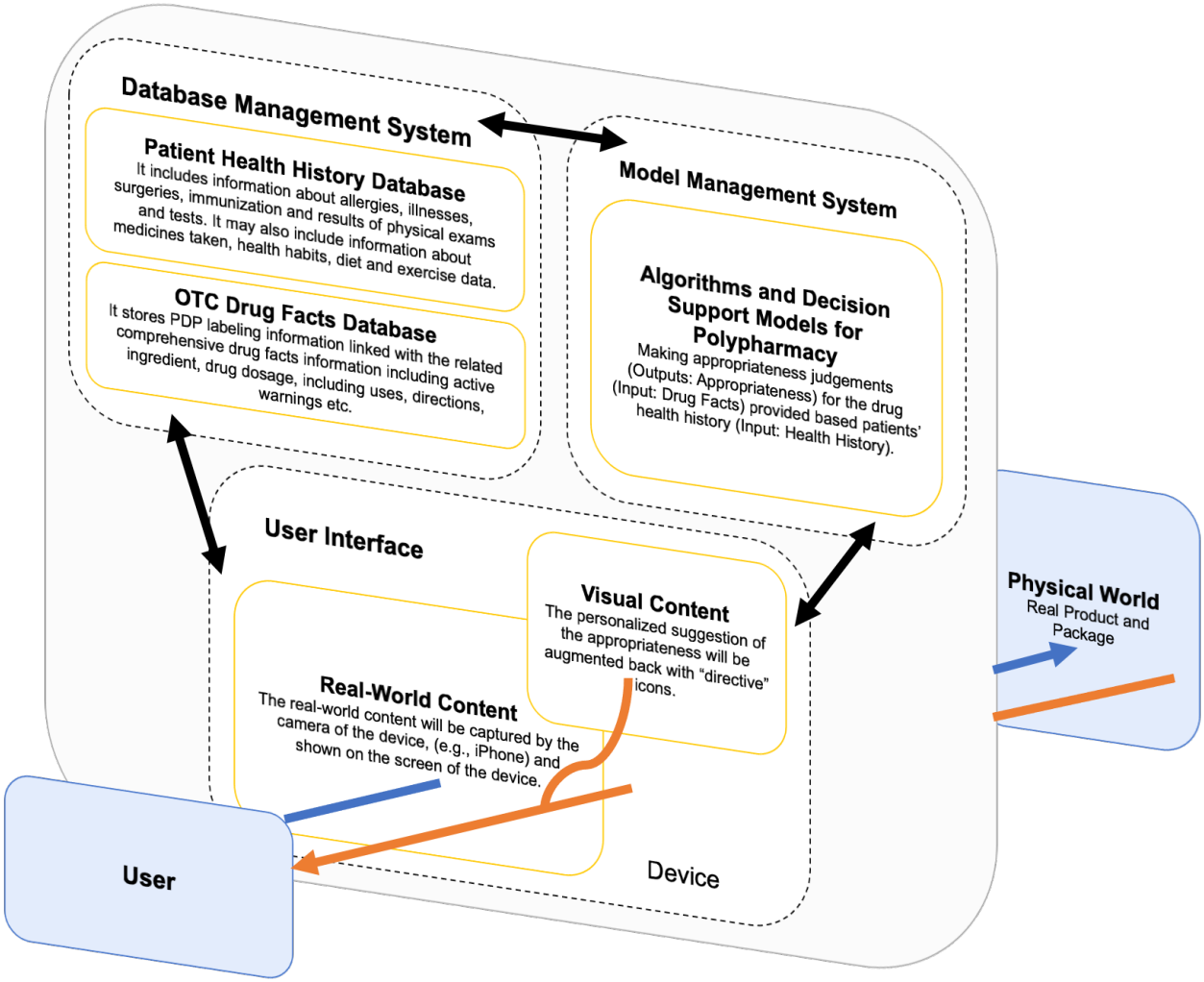


Figure 3.6 Framework of Self-medication Decision Support System with Personalized Labeling Concepts

The data management system included two databases: patient's health history database and the OTC drug facts database. The patient's health history database manages information about health conditions and medication history. It could also include information about health habits, diet, and exercise data. And the OTC drug facts database stored PDP labeling information linked with the related comprehensive drug facts information including active ingredient, drug dosage, uses, directions, warnings etc.

The model management system should contain algorithms or rules in decision support models related polypharmacy and contraindications. It took inputs from the DFL as well as the

patients' health history and make appropriateness judgement following rules or algorithms as output.

3.4 User Interface Concept for Personalized FOP Labeling via Augmented Reality

Based on the system framework above, a flowchart of the user interface concept for the personalized FOP labeling strategy was developed in Figure 3.7. Specifically, the workflow of the program started with the real-world content (e.g., package), which could be captured by a smart device (e.g., iPhone or iPad) by consumers. The principal display panels of OTC packages could be recognized and handled by the application on the smart device. Then, the smart device which connected to the database management system and the model management system could conduct logical comparisons and return results, which could be in the range of "appropriate to use" (green check mark), "do not use" (red stop sign) or "no answer but warnings needed to be noticed" (yellow alert sign). Last but not the least, the smart device could augment the relevant sign back to the principal display panel in a proper presenting way.

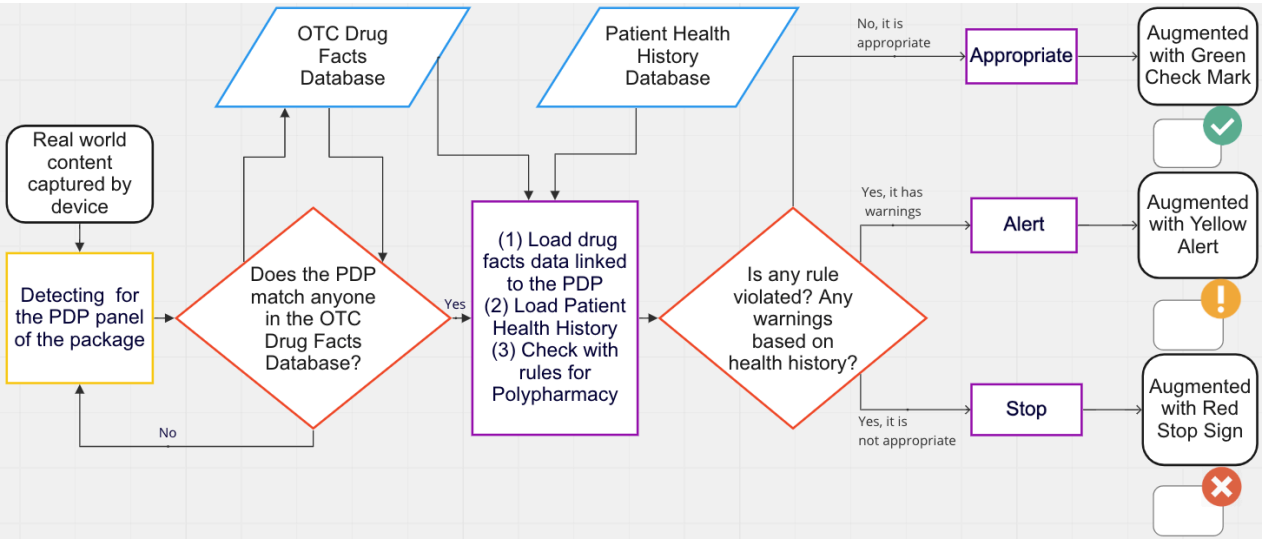


Figure 3.7 Flowchart of personalized labeling concept with data-driven and augmented reality

3.4.1 Materials and Methods

The augmented reality prototype of the personalized labeling concept for this aim was built using Xcode version 12.4 (12D4e) (Apple Inc, California, USA) under the Mac OS Catalina version 10.15.7. The testing devices were an iPhone 11 and iPad Pro 10.5-inch version. Images and graphics were developed with Adobe Creative Suite 2020 (Photoshop, Illustrator); packaging prototypes were developed by Esko ArtiosCAD v14. The prototype layouts can be found in APPENDIX E.3.

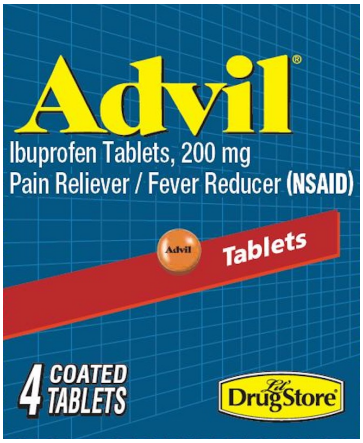
3.4.2 Software Development Kits

To realize the functions of user interface, image processing and augmented reality in this personalized labeling application, the UIKit, SceneKit and ARKit were the main software development kits this program was used. The reference links of the detailed documentations of those three kits were listed in the APPENDIX E.2. Briefly, the UIKit framework provides the required infrastructure for an iOS application; the SceneKit combines a high-performance rendering engine to handle 3D rendering objects; as well as the ARKit produces augmented reality by integrating iOS device camera and motion features.

3.4.3 Development of Personalized Labeling iOS Augmented Reality Application

To reiterate the objectives of this Aim, our goal of this development was to provide a working prototype of the user interface concept of personalized labeling strategy with augmented reality. Our focuses were on the development of user interface rather than the database management system nor the model management system. As such, we used four real-market OTC packages as design samples to simplify database (Figure 3.8 and APPENDIX E.3). Also, the returned results of appropriateness were also pre-defined as follows based on a fictitious patient as an avatar: (1) NyQuil – “appropriate to use” with green check mark; (2) Advil – “do not use”

with red stop sign; (3) Tylenol and TopCare pain reliever – “no answer but warnings needed to be noticed.” with yellow alert sign. Figures 3.7 to 3.10.



(3) PDP of Advil

(1) PDP of Tylenol PM

(2) PDP of TopCare Pain Relief PM

(4) PDP of NyQuil

Figure 3.8 Four Real-market OTCs as Prototyping Samples.

For the details of the user interface program, the source codes and installation instructions were uploaded via the GitHub repository. APPENDIX E For the program, a simple camera view user interface was generated via UIKit. The iPhone camera was activated as default when opening the application, and it started to search for any physical planes in the real-world content which matched the reference images with the help from ARKit functions. The PDPs of the four OTC samples were saved as reference images for the camera to search. Once the camera found an object in its view matched one of those reference images, it would immediately lock the object and set it as an AR image anchor. Based on the AR image anchor and the pre-defined returned results of appropriateness for the reference image, we could augment animated flashing symbols, textboxes back to the detected plane in the real-world at real time, with the help of the

functions from SceneKit and ARKit. Figure 3.9 shows the pictures of user interface prototype of personalized labeling application.

Figure 3.9 Personalized Labeling Prototyping

Chapter 4 Evaluating the Potential Effectiveness of the Personalized Labeling Concept as a Support for OTC Decision Making (Aim 3)

4.1 Objectives

Aim 3 focuses efforts on evaluating the potential benefits that a personalized, FOP labeling strategy would provide in terms of its effectiveness (ability to make the correct decision) and efficiency (time to correct decision) when people evaluate an OTC for use given a specific scenario. Results of the novel approach will be compared with the performance of the current, commercial approach to labeling.

Specifically, we hypothesized that consumers who were educated regarding the novel label system would spend less time making more appropriate decisions when the personalized labeling systems were applied compared to trials which employed the existing, commercial approach. Herein, we provide preliminary data regarding the potential benefit of such a labeling strategy.

4.2 Method

4.2.1 Overview of Experiment Design

In support of the objectives, we used an absolute judgment method utilizing a custom program built in E-prime 3.0. (*Tools Psychology Software, Pennsylvania, USA*). Stimuli were created with the resolution of 1920x1080 using Adobe Illustrator 2020. The testing program was run on a Dell laptop workstation with E-run 3.0. (Intel Core i5-7440HQ CPU, 16GB RAM, 238GB memory, a 13-inch 1920*1080 display and 64bit Windows 10 operating system).

Participants were assigned into two groups. In the “concept-educated” group participants were introduced to the concept of personalized labeling. Specifically, they were informed of the meaning of the green “checkmark” symbol and the red “stop sign” symbol and that this would

test the efficacy of an application that utilized augmented reality, returning a customized response regarding appropriateness for those who utilized the application. Participants in the control group were not educated about the personalized labeling concept. (APPENDIX F)

Each participant in both groups of this study completed a total of 44 absolute judgement trials which posed the question of drug appropriateness for a given scenario posed by the question. Trials were organized as shown in Figure 4.1. Participants began with four practice trials intended to acquaint them with the test procedure and enable them to clarify any questions that they had with the research team prior to beginning testing. The trials of interest, or test trials, were comprised of 32 absolute judgement tasks which were divided into four blocks needed to accomplish our randomization scheme. Upon completing each block of test trials, participants were provided two "dummy trials" intended to minimize any possible order effects resulting from short term memory. Practice trials and dummy trials were not included in the analysis.

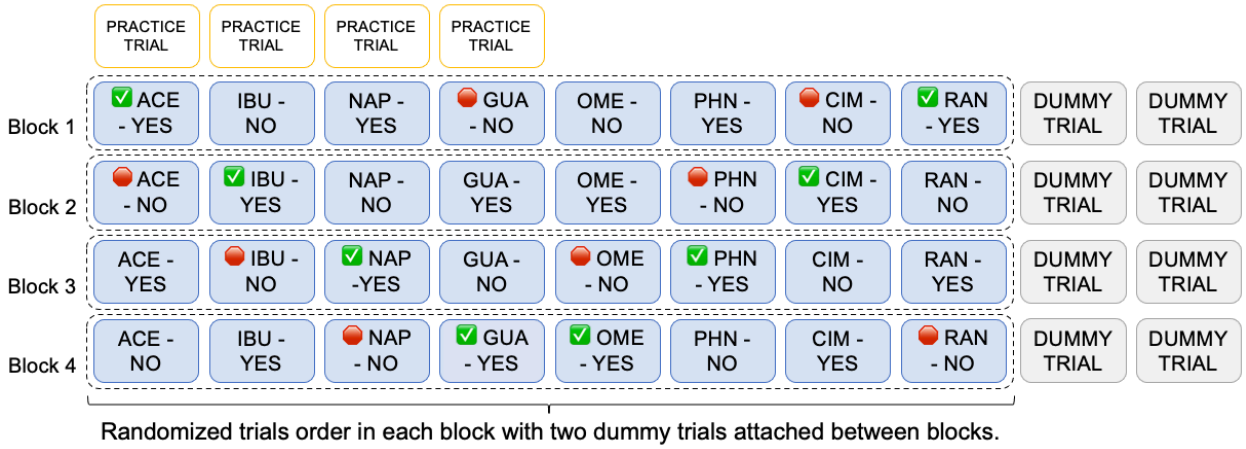


Figure 4.1 Trials of the Absolute Judgement Test

Test trials were counterbalanced regarding "correct response" (i.e., "Yes, this is appropriate" given the scenario case vs "No, it is not"). Half of the participants were educated regarding the personalized FOP labeling concept in advance of the experiment and half were not.

Stimulus were based on single-ingredient, real-world commercial brands. The following active ingredients were tested: acetaminophen (ACE) – Tylenol; ibuprofen (IBU) – Advil; naproxen (NAP) – Aleve; guaifenesin (GUA) - Mucinex, omeprazole (OME) – Prilosec; phenylephrine (PHN) – Sudafed; cimetidine (CIM) – Tagamet; ranitidine (RAN) – Zantac. Answer was offered in a binary form (Appropriate for use yes/no) and crossed with design at two levels (personalized FOP and standard), for a total of four treatments. An example of Tylenol (Acetaminophen) in the personalized FOP presenting a scenario question which results in an "appropriate to use" response is shown in Figure 4.2, and the entire trial images for the absolute judgement test are shown in APPENDIX G.

Figure 4.2 An Example of Tylenol with Check Mark Symbol Augmented and the Scenario Question with "Appropriate to Use" Answer

To minimize any possible order effects, a stratified randomization scheme was used for trials. This scheme divided trials into 4 blocks as shown in Figure 4.1. One of the four treatments (correct response-appropriate/inappropriate x label design-personalized

FOP/standard) of each brand/active ingredient were randomly assigned to a single block. In doing so, any brand/active ingredient only appeared once per block. Within each block, run order was randomized. Between blocks, two dummy trials were added to minimize the practice effects of short-term memory carry over, minimizing the likelihood that participants would see the same active ingredient in back-to-back trials.

The task for the 32 test trials was comprised of an absolute judgment (appropriateness-yes/no) based on scenarios drafted by the research team that could be answered with information directly available from the Drug Facts Label (DFL). Of the test trials, a total of 16 questions had a "yes" appropriate response and 16 were not appropriate for the person in the scenario to take. Appropriate and inappropriate correct answers were balanced across brands and design (standard vs personalized FOP).

As an example of Tylenol shown above in Figure 4.2, question design was constructed using the following rules: (1) Each trial question was constructed to be no longer than two lines in length when using 18-point Arial font. (2) The sentence asked whether the product depicted in the trial was appropriate for the person/avatar to use based on a scenario, e.g. "*Is this pain reliever appropriate to use for a person taking melatonin?*" Firstly, if the question for the affirmative response, "appropriate to use" (YES), then the scenario within the question was not related to information provided using the Drug Facts Label (DFL). By contrast, if the question is designed to have an answer "not appropriate to use" (NO), then the scenarios were drafted to be a direct match to warning information from the DFL (i.e., overdose of active ingredient, drug-drug interaction, or drug-diagnosis interaction).

4.2.2 Participants and Recruitment

Participants were recruited via the SONA recruiting system available from the MSU College of Communication Arts and Sciences, distribution of IRB approved flyers, and word of mouth. (APPENDIX H) Participants were eligible for this study if they met all the following criteria:

- Were at least 18 years old
- Not legally blind
- Had used OTC drugs during the 6 months preceding the experiment
- Had the ability to come to the Healthcare, Universal Design, Biomechanics lab (HUB) at Michigan State University for the test.

4.2.3 Procedure

Upon arrival at the lab, where applicable, participants were provided a 1.5-hour parking validation. Prior to the experiment, they were asked to read and sign the IRB-approved consent form (APPENDIX I) and were verbally informed of their right to stop completely or opt out of any portion of the testing, and that they will still the $25 cash incentive in these cases. Data were recorded by participant number, with no link to participants' identities.

4.2.2.1 Pre-tests

Once informed, written consent was obtained, participants were characterized using a brief survey of demographics; this was followed by a series of tests which characterized participants, referred to henceforth as pre-tests. (APPENDIX D)

The pre-tests include three standard tests. Detailed procedures relating to the same are located in APPENDIX D. The standard testing methods were mentioned in the change detection

chapter (Chapter 2.4.1), which included: (1) Near-point visual acuity test; (2) The Rapid Estimate of Adult Literacy in Medicine, Revised (REALM-R); (3) Color differentiation test.

4.2.2.2 Introduction of Personalized Labeling Concept

For the "concept-educated" group, participants were guided to a table where several packaging prototypes and an iPhone 11 were placed. The researcher introduced the personalized labeling concept by reading the script depicted in the APPENDIX F. During the introduction, participants were guided to interact with the personalized-labeling-augmented packaging prototypes with the iPhone 11 which installed the personalized labeling application as shown in the Figure 4.3. It should be noted that we improved the language in the instruction to convey the ideas of personalized labeling strategy to be more straightforward than older version of the script. Specifically, we directed participants to the green check mark which indicated "appropriate-to-use" and the stop sign which indicated "not-appropriate-to-use". Also, upon conclusion of the introduction, the following explicit instruction was also added in the new version of the instruction.

> *"In the following test, half of the trials you are about to undertake assume that the person in the scenario is using the app and it is returning a response that is specific to the drug and its appropriateness for the person. Half of the trials assume that the person is not using this app. Please answer the questions for the person in the trial as quickly as you can."*

The reasons for deciding to update the instruction are discussed in the results section. The old and new versions of the instruction documents are listed in the APPENDIX F.

Figure 4.3 The Introduction of Personalized-Labeling Concept to Participants in the Concept-Educated Group. (a) The prototypes were on the lab desk; (b) The prototypes were viewed via Personalized-Labeling iOS application; (c) A prototype was picked up for closer view by participant.

4.2.2.3 Absolute judgement test

After the pretests. and introductory educational session for applicable participants, participants were comfortably seated at Dell laptop which ran the absolute judgement testing program using E-prime 3.0 software. Researchers coded the pretest data prior to running the test program. Coded information included: participant number, computer number, subject's sex, age, ethnicity, educational level, native language, near point visual acuity score, REALM-R health literacy score, and color differentiation ability.

The absolute judgement task began with a review of screening criteria to reconfirm eligibility. Once participants confirmed met all eligibility criteria, and reconfirmed that they had no history of seizure, a brief introduction of the experiment appeared on the screen to welcome each participant and instruct them regarding the testing details and how to begin the program.

"Welcome to the experiment!"

"During this test, we will show you some medication images. Above each medication image, we ask one question with given scenarios. Please answer the question for the person in each scenario. As soon as you decide, press "1" key for yes, and "0" key for no. The program will direct you to the next image, and so forth, until to the end of this test.

Now, let's start with a few practice trials. Please press SPACEBAR to continue once you are ready."

After the introduction page, participants were directed to finish four practice trials. The objective of the practice trials (see APPENDIX G) was to familiarize the participants with the program and provide them with an opportunity to ask questions of the research team prior to beginning the experiment. Once a participant completed the four sample trials, instructions were presented related to the test trials.

"INSTRUCTIONS"

"Now, you will have the main test. In each trial of this test, we will ask you to answer a question."

"Please assume that, in each trial, the person in the scenario has the conditions/symptoms which the drug treats, and then answer whether or not the pictured drug is appropriate for the person to take."

"As soon as you decide, press the '1' key for 'Yes, it is appropriate to use', and the '0' key for 'No, it is not appropriate'."

"If you have any questions, please ask now, since researchers will not answer any questions once the trials start. When you are ready, please press the SPACEBAR to start."

4.2.4 Post-test Debriefing for Manipulation Check

After the main test, participants were guided to a private room for the post-test debriefing session. They were asked the debriefing questions presented in APPENDIX J. After the debriefing session, participants were thanked and compensated with $25 for their participation.

4.3 Power Calculations

Power estimates for this study were based on previous work (L. Bix et al., 2016) suggesting an effect size of d=0.84. The previous work focused on surgical technicians and healthcare providers who were required to select appropriate medical devices as part of their job (surgical technologists), the power estimate was conducted with 30% of the measured effect size for this study which employed a more general population, or d=0.25. For this within subject study with 8 repetitions, 66 participants (recruiting 82 participants before attrition) was estimated as allowing to detect d=0.25 with given α=0.05 and power >0.8.

4.4 Variables and Measurements

To test the effectiveness and efficiency related to the use of a personalized FOP labeling strategy on decision making, we focused on two dependent variables: (1). the accuracy of participants' responses (a binary variable, yes or no), representing the design's effectiveness; (2) the time to correctly answer the question (a continuous variable, in seconds) representing a measure of the efficiency. Data were processed and tested for normality with appropriate transformations (details were included in Chapter 4.5.2 and Chapter 4.5.3 respectively)

The predictor variables included in the final analysis were: the between-subject variable-education of the personalized labeling concept (the concept-educated group vs the control group), abbreviated as "**PerLab**"; Design **Layout** (personalized FOP vs. standard); **Question Type** (questions with "Yes, appropriate" as correct answer vs. questions with "No, not appropriate" as correct answer); Ingredients (8 drug active ingredients); education (above bachelor degree vs. bachelor or below); ethnicity (white or other), sex (male or female); language (English or other) and age.

4.5 Results

4.5.1 Descriptive Statistics

A total of 86 participants were recruited via the paid MSU SONA system for this study in the late summer and fall of 2021 at the Lab of Healthcare, Universal design, and Biomechanics in the School of Packaging. Seventy-two participants were included in the final analysis. Fourteen study participants were not included in the analysis. These early participants were part of the concept-educated group. Specifically, they received instructions that did not seem to clearly communicate the new paradigm of a personalized labeling concept. As a result of their failure to understand the concept, the research team created a new version of the concept education instructions; specifically, we attempted to better communicate the tool and also create a stronger link to the experiment that the participants were about to undertake. This change was catalyzed by data we saw in early debriefing sessions, where a majority, 71.4%, 10 out of 14, indicated that they had not noticed or utilized the personalized system in their decision-making process. From the 15th participant forward in the concept-educated group, we used the new concept-education instruction, as well as excluded those participants who received the initial instruction script. The detailed changes of instructions were mentioned previously in the Method Chapter 4.2.2.2 and the two versions of instruction can be found in the APPENDIX F.

Table 4.1 provides frequencies of participation across demographic factors of interest for the entire study population. Among the 72 participants included data set that was analyzed, half (N1=36) were assigned to the concept educated group and half to the control group (N0=36). The final sample for analysis included 8 men and 28 women with an average age of 34.64 (SD=13.23) were in the concept educated group, while 11 men and 25 women with an average

age 34.19 (SD=14.90) were in the control group. The detailed age distributions in each group are shown in the Figure 4.4-4.5.

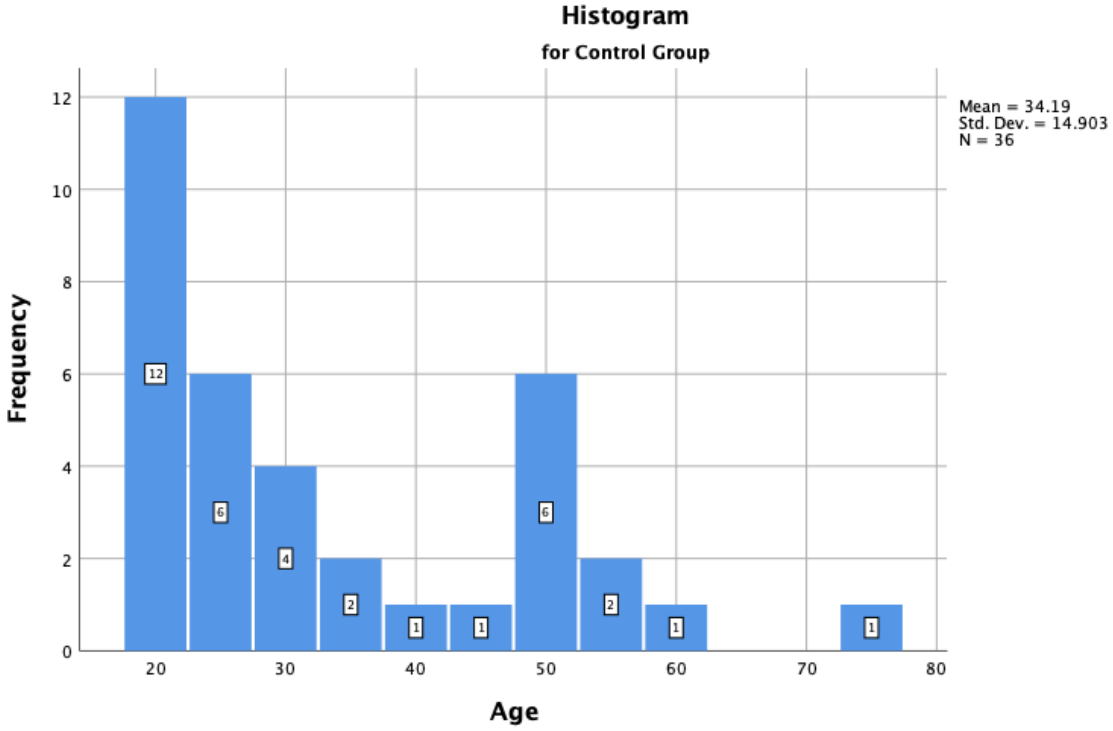


Figure 4.4 Histogram of Age for the Control Group

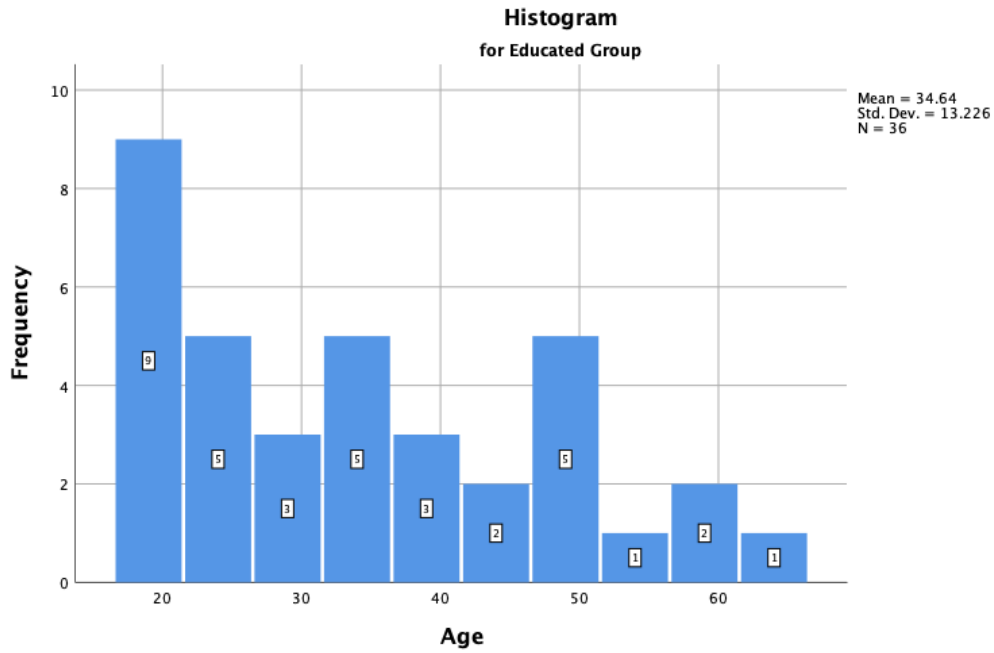


Figure 4.5 Histogram of Age for the Educated Group

Table 4.1 Frequencies of Participation across Demographic Factors of Interest for the Entire Study Population

Characteristic	Group	Value (72)	Number	% of Total
Sample size	**Educated**	N1 = 36 (50%)		
	Control	N0 = 36 (50%)		
Sex	**Educated**	Male	8	11.1%
		Female	28	38.9%
	Control	Male	11	15.3%
		Female	25	34.7%
Color Differentiation	**Educated**	Normal	36	50.0%
		Color blind	0	0.0%
	Control	Normal	35	48.6%
		Color blind	1	1.4%
Education	**Educated**	Masters or higher	14	19.4%
		Bachelor or lower	22	30.6%
	Control	Masters or higher	8	11.1%
		Bachelor or lower	28	38.9%
Ethnicity	**Educated**	White	22	30.6%
		Others	14	19.4%
	Control	White	24	33.3%
		Others	12	16.7%
Health Literacy	**Educated**	Normal	35	48.6%
		Risk for poor literacy	1	1.4%
	Control	Normal	35	48.6%
		Risk for poor literacy	1	1.4%
Native Language	**Educated**	English	27	37.5%
		Others	9	12.5%
	Control	English	30	41.7%
		Others	6	8.3%
Visual Acuity	**Educated**	Normal (<=20/40)	36	50.0%
		Poor (>20/40)	0	0.0%
	Control	Normal (<=20/40)	36	50.0%
		Poor (>20/40)	0	0.0%
Age	**Educated**	Mean (Min, Max)	34.64	Min = 19, Max = 65
		Std. Deviation	13.226	
	Control	Mean (Min, Max)	34.19	Min = 20, Max = 73
		Std. Deviation	14.903	

4.5.2 Response Accuracy

A generalized linear mixed model was fitted to assess the influence of the variables of interest on response accuracy, or the probability of correctly answering the question for a given trial scenario. The binary data - correctly answered questions or not (yes/no) were interpreted in terms of probability (p) of correctly answering trial questions with logit transformation, *ln (p/1-p)*; and the results of estimated means were back-transformed and displayed in terms of original target scale.

The predictor variables included in the final model, as aforementioned in Chapter 4.4, were "PerLab", layout, question type, ingredients, education, ethnicity, sex, language and age. All possible 2-way and 3-way interactions among PerLab, layout and question type were also included. The participants themselves were included as random effects. All estimated means were back transformed to the original scale of the dependent variable in percentage.

A total of 2,304 trials (72 participants x 32 trials) were analyzed in this absolute judgement test to examine data for effects on the dependent variable response accuracy. Participants provided answers correctly in 2,141 trials (92.9%), and incorrectly responded in 163 trials (7.1%).

A summation of the results from the statistical analysis is presented in Table 4.2. For the variables of interest, there were significant fixed effects of design layout (p=0.002) on response accuracy. A significant main effect of Layout was indicated (p=0.002), with the presence of a personalized FOP labeling enhancing participant response accuracy. While there was no evidence of main effects associated with concept education, "PerLab" (p=0.268), or question types (p=0.468) on the response accuracy, there was a significant 2-way (p=0.003) when the concept education "PerLab" and design layout were crossed. Not surprisingly, the presence of an

FOP personalized labeling significantly improved the accuracy for the concept educated group as compared to their performance with the standard label but was not found to significantly improve performance for the control group, who had not been informed of its purpose.

For the covariates, the impact of ingredients was found statistically significant ($p<0.001$) on response accuracy. Beyond that, no other covariates or 2-way or 3-way interactions between variables of interest were found to have statistically significant impacts on response accuracy. (Table 4.2)

Table 4.2 Tests of Fixed Effects on Response Accuracy (bolded effects are significant at alpha = 0.05.)

Effects	*F*	*df1*	*df2*	*Sig.*
Corrected Model	3.036	19	2284	0.000
PerLab	1.227	1	2284	0.268
Design Layout	**9.553**	**1**	2284	**0.002**
QuestionType	0.526	1	2284	0.468
Ingredient	**4.985**	**7**	2284	**0.000**
Education	0.243	1	2284	0.622
Ethnicity	0.433	1	2284	0.510
Sex	0.222	1	2284	0.638
Language	2.174	1	2284	0.140
Age	0.012	1	2284	0.912
PerLab * Layout	**8.567**	**1**	2284	**0.003**
*PerLab * QuestionType*	0.820	1	2284	0.365
*Layout * QuestionType*	0.018	1	2284	0.893
*PerLab * Layout * QuestionType*	0.089	1	2284	0.765

Probability distribution: Binomial ; Link function: Logit

4.5.2.1 Effect of Layout

To explore the main effect of design layout noted above, results of estimated means of layout and simple contrasts between layout levels (personalized FOP vs standard) are presented in the Tables 4.3-4.4. Specifically, there was evidence for statistically improved response accuracy ($p=0.006$, contrast estimate=0.025) when the personalized FOP layouts (ME=0.964, SE=0.008) as compared to response accuracy for trails comprised of standard layouts

(ME=0.939, SE=0.013). To conclude, the participants have higher response accuracy when personalized FOP are shown. Beyond that, since a significant 2-way interaction between PerLab (concept education) and layout was found, more details of the effects of layout within the interaction will be discussed in the following section.

Table 4.3 Results of Estimated Means of Layout

Layout	*Mean*	*Std. Error*	*95% Confidence Interval*	
			Lower	Upper
Personalized Front-of-Pack	0.964	0.008	0.944	0.978
Standard	0.939	0.013	0.908	0.96

Table 4.4 Simple Contrasts of Layout

Layout Simple Contrasts	*Contrast Estimate*	*Std. Error*	*t*	*df*	*Adj. Sig.*	*95% Confidence Interval*	
						Lower	Upper
Personalized Front-of-Pack vs. Standard	0.025	0.009	2.745	2284	0.006	0.007	0.044

The sequential Bonferroni adjusted significance level is .05.
Confidence interval bounds are approximate.

4.5.2.2 Significant 2-way interaction: PerLab x Layout

To test the hypothesis of this study, pairwise comparisons were conducted to interpret the significant 2-way interaction that was identified between the concept-educated groups (those informed about the personalized labeling strategy and those not (control group) - PerLab) and design layout (personalized labeling/standard). When the personalized labeling concept was introduced prior to the test (Table 3.5), participants were significantly more likely to answer correctly in the trials with personalized FOP layouts (ME=0.977, SE=0.007) as compared to the trials comprised of the standard formatting (ME=0.933, SE=0.017) (p=0.002). In contrast, for the control group (Table 3.5), there was no evidence of a significant difference of response accuracy when trials with personalized FOP layouts (ME=0.946, SE=0.015) were compared to standard formats (ME=0.944, SE=0.015) (p=0.898). From another perspective, when the personalized

FOP was depicted (Table 3.4), people who were introduced to the personalized labeling concept (ME=0.977, SE=0.007) were significantly more accurate in their question response compared to the control group (ME=0.946, SE=0.015), who had not been acquainted to the concept prior to the study (p=0.002). In contrast, for the standard commercial layouts, no evidence of significant difference was evident when the response accuracy of those introduced to the concept of personalized labeling (ME=0.933, SE=0.017) was compared to those in the control group (ME=0.944, SE=0.015). (p=0.574)

To test the hypothesis of this study, pairwise comparisons were conducted to interpret the significant 2-way interaction that was identified between the concept-educated groups (those informed about the personalized labeling strategy and those not - PerLab) and design layout (personalized labeling/standard). (Figure 4.6 and Tables 4.5-4.7)

Table 4.5 Results of Estimated Means of the Interaction between PerLab and Layout

PerLab	*Layout*	*Mean*	*Std. Error*	*95% Wald Confidence Interval*	
				Lower	Upper
Concept-educated Group	Personalized Front-of-Pack	0.977	0.007	0.957	0.988
	Standard	0.933	0.017	0.891	0.960
Control Group	Personalized Front-of-Pack	0.946	0.015	0.908	0.969
	Standard	0.944	0.015	0.906	0.968

Continuous predictors are fixed at the following values: Age=34.42

When the personalized labeling concept was introduced prior to the test (Table 4.6), participants were significantly more likely to answer correctly in the trials with personalized FOP layouts (ME=0.977, SE=0.007) as compared to the trials comprised of the standard formatting (ME=0.933, SE=0.017). (p=0.002) In contrast, for the control group (Table 4.6), there was no evidence of a significant difference of response accuracy when trials with personalized FOP layouts (ME=0.946, SE=0.015) were compared to standard formats (ME=0.944,

SE=0.015). (p=0.898) From another perspective, when the personalized FOP was depicted (Table 4.7), people who were introduced to the personalized labeling concept (ME=0.977, SE=0.007) were significantly more accurate in their question response compared to the control group (ME=0.946, SE=0.015), who had not been acquainted to the concept prior to the study (p=0.002). In contrast, for the standard commercial layouts, no evidence of significant difference was present was found when the response accuracy of those introduced to the concept of personalized labeling (ME=0.933, SE=0.017) was compared to those in the control group. (ME=0.944, SE=0.015) (p=0.574)

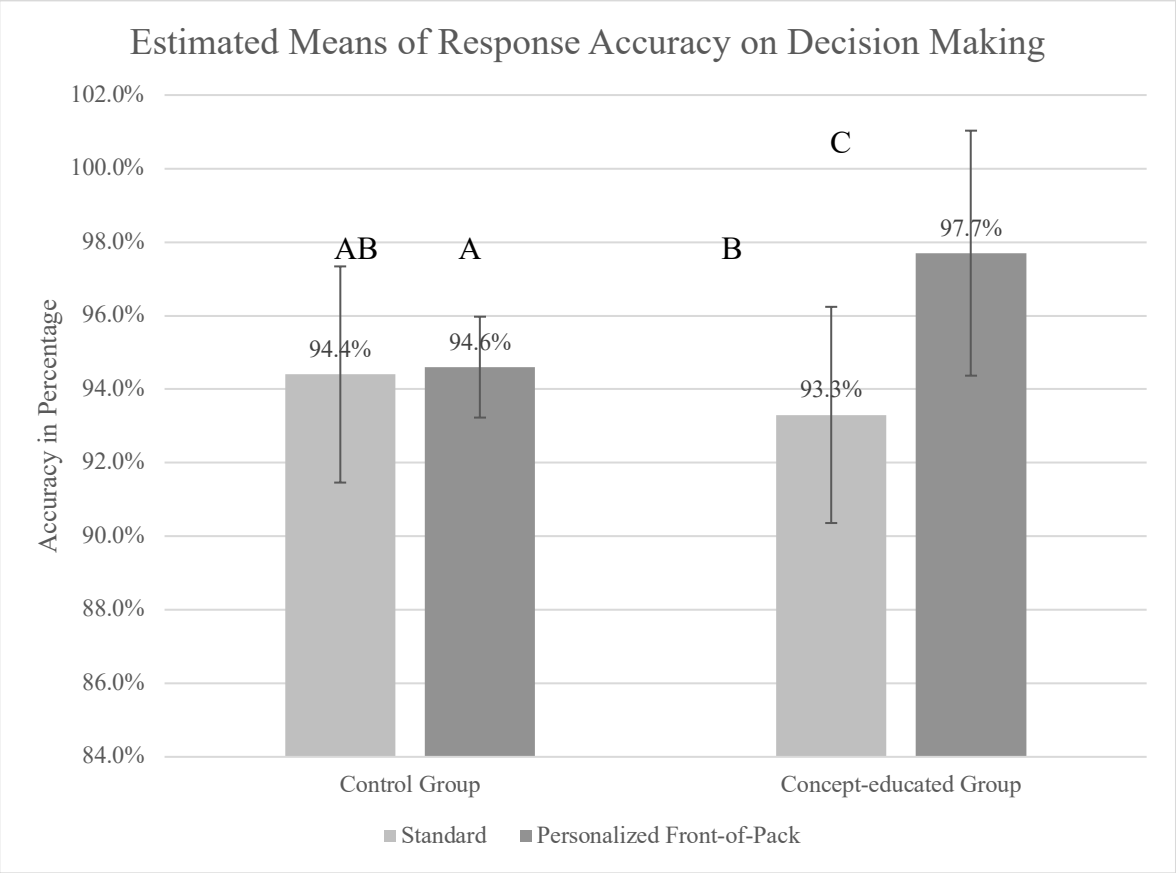


Figure 4.6 Estimated Means of Response Accuracy on Decision Making

Table 4.6 Pairwise Comparisons of Effect of Layout on Response Accuracy under each Perlab level

PerLab	*Layout Pairwise Contrasts*		*Contrast Estimate*	*Std. Error*	*t*	*df*	*Adj. Sig.*	*95% Confidence Interval*	
								Lower	Upper
Concept-educated Group	Personalized Front-of-Pack	Standard	0.043	0.014	3.088	2284	**0.002**	0.016	0.071
	Standard	Personalized Front-of-Pack	-0.043	0.014	-3.088	2284	**0.002**	-0.071	-0.016
Control Group	Personalized Front-of-Pack	Standard	0.002	0.012	0.128	2284	0.898	-0.022	0.025
	Standard	Personalized Front-of-Pack	-0.002	0.012	-0.128	2284	0.898	-0.025	0.022

The sequential Bonferroni adjusted significance level is .05; Confidence interval bounds are approximate.

Table 4.7 Pairwise Comparisons of Effect of PerLab on Response Accuracy under each Layout level

Layout	*PerLab Pairwise Contrasts*		*Contrast Estimate*	*Std. Error*	*t*	*df*	*Adj. Sig.*	*95% Confidence Interval*	
								Lower	Upper
Personalized Front-of-Pack	Concept-educated Group	Control Group	0.031	0.015	2.064	2284	**0.039**	0.002	0.061
	Control Group	Concept-educated Group	-0.031	0.015	-2.064	2284	**0.039**	-0.061	-0.002
Standard	Concept-educated Group	Control Group	-0.011	0.019	-0.562	2284	0.574	-0.049	0.027
	Control Group -	Concept-educated Group	0.011	0.019	0.562	2284	0.574	-0.027	0.049

The sequential Bonferroni adjusted significance level is .05; Confidence interval bounds are approximate.

4.5.2.3 Effect of Ingredients

The ingredient being tested in the trial was also found to affect the accuracy of participant response statistically significantly. Results related to this finding are presented in the Tables 4.8-4.9. Simple contrasts utilize the response to trials containing acetaminophen as a base.

Table 4.8 Results of Estimated Means of Ingredients

Ingredient	*Mean*	*Std. Error*	*95% Confidence Interval*	
			Lower	Upper
Ranitidine (acid reducer)	0.909	0.023	0.854	0.944
Phenylephrine (nasal decongestant)	0.958	0.013	0.922	0.977
Omeprazole (acid reducer)	0.948	0.015	0.909	0.971
Naproxen (pain reliever)	0.973	0.01	0.944	0.987
Ibuprofen (pain reliever)	0.916	0.021	0.863	0.949
Guaifenesin (cough suppressant)	0.995	0.004	0.978	0.999
Cimetidine (acid reducer)	0.909	0.023	0.854	0.944
Acetaminophen (pain reliever)	0.909	0.023	0.854	0.944

Continuous predictors are fixed at the following values: Age=34.42

Table 4.9 Simple Contrasts of Ingredients with Acetaminophen as Basis

Ingredient Simple Contrasts	*Contrast Estimate*	*Std. Error*	*t*	*df*	*Adj. Sig.*	*95% Confidence Interval*	
	Acetaminophen (pain reliever) as comparing base					Lower	Upper
Ranitidine (acid reducer)	-1.33E-15	0.024	-5.57E-14	2284	1	-0.047	0.047
Phenylephrine (nasal decongestant)	0.049	0.022	2.266	2284	0.118	-0.007	0.105
Omeprazole (acid reducer)	0.04	0.022	1.817	2284	0.278	-0.015	0.094
Naproxen (pain reliever)	0.064	0.022	2.967	2284	0.018	0.007	0.121
Ibuprofen (pain reliever)	0.007	0.023	0.292	2284	1	-0.049	0.063
Guaifenesin (cough suppressant)	0.086	0.022	3.876	2284	0.001	0.026	0.146
Cimetidine (acid reducer)	-1.33E-15	0.024	-5.57E-14	2284	1	-0.047	0.047

The sequential Bonferroni adjusted significance level is .05. Confidence interval bounds are approximate.

The results of comparisons suggest that people were significantly more accurate for the trials that asked questions about guaifenesin (ME=0.995, SE=0.023; p=0.001), naproxen (ME=0.973, SE=0.010; p=0.018) as compared to the trials involving responses associated with acetaminophen (ME=0.909, SE=0.023). No other ingredients were found significantly different on response accuracy comparing to Acetaminophen.

4.5.3 Response Time

In addition to the dependent variable accuracy of response, we also evaluated time to correct response as a dependent variable - response time (in seconds for those responses that were correctly answered). The response time data set were checked for the validity of normality assumptions prior to statistical analyses. Residual plots and normal probability plots of the original data suggested an appropriate transformation was needed. As a result, data were natural log transformed. Tukey's method was used for minor non-constant variance and Satterthwaite's degree of freedom was used to adjust degrees of freedom. A generalized linear mixed model was then fitted to this natural log-transformed response time variable. A total of 2,141 correctly answered absolute judgement test trials were included in the data analysis

The predictor variables included in the final model were the same as the previous analysis, namely: PerLab, design layout, question type, ingredients, education, ethnicity, sex, language, and age. All possible 2-way, 3-way interactions across PerLab, layout and question type were also included. Additionally, the following covariates were included in the final model: Participants themselves were also included as random effects. All estimated means were back transformed to the original scale of the dependent variable in seconds.

A summation of analysis results is presented in Table 4.10. Evidence was found to suggest significance for the main, fixed effects of concept education (PerLab) ($p<0.001$); design

layout (personalized labeling vs standard labeling) ($p<0.001$); question type (appropriate for use under a given scenario- yes/no) ($p<0.001$) and ingredients ($p<0.001$). Additionally, all 2-way interactions were found to have statistically significant impacts on the time to correctly respond: specifically, PerLab x layout ($p<0.001$), PerLab x question type ($p=0.000$) and layout x questions type ($p<0.001$). The 3-way interaction was also found statistically significant; concept education (Perlab) x layout x question type ($p=0.015$). More details are presented in the following subsections.

Table 4.10 Tests of Model Effects on Response Time for Correct Responses

Source	*F*	*df1*	*df2*	*Sig.*
Corrected Model	29.035	19	204	0.000
PerLab	**21.629**	**1**	**65**	**0.000**
Layout	**111.276**	**1**	**2056**	**0.000**
QuestionType	**172.187**	**1**	**2057**	**0.000**
Ingredient	**15.819**	**7**	**2057**	**0.000**
Education	3.8	1	65	0.056
Ethnicity	0.811	1	65	0.371
Language	2.711	1	65	0.105
Sex	0.13	1	65	0.719
Age	0.263	1	65	0.610
PerLab * QuestionType	**6.769**	**1**	**2057**	**0.009**
PerLab * Layout	**107.43**	**1**	**2056**	**0.000**
QuestionType * Layout	**12.398**	**1**	**2057**	**0.000**
PerLab * QuestionType * Layout	**5.928**	**1**	**2057**	**0.015**

4.5.3.1 Effects of PerLab, Layout, Question Type, Ingredients

Table 4.11-4.12 present statistical results of estimated marginal means and simple contrasts for the significant fixed effects of variables as aforementioned: PerLab, layout, question type and ingredients. The estimated marginal means were back transformed to the original scale in seconds.

Table 4.11 Estimated Marginal Means of PerLab, Layout, Question Type

	Mean	*Std. Error*	*95% Confidence Interval*	
			Lower	Upper
PerLab				
Concept-educated Group	13.513	1.166	11.373	16.055
Control Group	22.167	1.992	18.526	26.523
Layout				
Personalized Front-of-Pack	14.553	1.048	12.607	16.8
Standard	20.581	1.486	17.823	23.766
QuestionType				
Question with "Yes" as correct answer	21.476	1.551	18.597	24.799
Question with "No" as correct answer	13.947	1.004	12.082	16.101
Ingredients				
Ranitidine (acid reducer)	19.736	1.639	16.744	23.262
Phenylephrine (nasal decongestant)	17.497	1.439	14.867	20.591
Omeprazole (acid reducer)	18.848	1.555	16.009	22.191
Naproxen (pain reliever)	17.245	1.414	14.66	20.286
Ibuprofen (pain reliever)	23.095	1.921	19.59	27.228
Guaifenesin (cough suppressant)	12.499	1.02	10.634	14.692
Cimetidine (acid reducer)	15.031	1.246	12.757	17.711
Acetaminophen (pain reliever)	16.528	1.375	14.02	19.485

For the simple contrasts within the levels are compared to a base and results of this comparisons are presented in Table 4.12, f the value of contrast estimate is positive, it means participants under such level for contrast took more time for correct decisions than trails related to the base; and if negative, then they took less time.

For the fixed effect of concept educations (PerLab), participants in the concept-educated group (ME=13.513, SE=1.166) took less time than those in the control group (ME=22.167, SE=1.992) with a difference of 8.654 seconds (SE=2.018, p=6.05E-05). That said, this main effect is impacted by significant interactions which were also identified and discussed below.

For the fixed effects of layout, for trails involving personalized FOPs (ME= 14.553, SE=1.048), participants took 6.028 seconds less time (SE=0.717, p=4.44E-16) to make correct decisions than when the standard labels were present (ME=20.581, SE=1.486). As with concept

education, the main effect was mediated by other factors as indicated and discussed in the section on interaction terms.

Table 4.12 Simple Contrasts of PerLab, Layout, Question Type

Simple Contrasts	*Contrast Estimate*	*Std. Error*	*t*	*df*	*Adj. Sig.*	*95% Confidence Interval*	
						Lower	Upper
PerLab	**(Control Group as basis)**						
Concept-educated Group	-8.654	2.018	-4.289	65	6.05E-05	-12.683	-4.625
Layout	**(Standard as basis)**						
Personalized Front-of-Pack	-6.028	0.717	-8.403	484	4.44E-16	-7.437	-4.618
QuestionType	**(Question with "No" as correct answer as basis)**						
Question with "Yes"	7.528	0.789	9.543	308	0.000	5.976	9.080
Ingredient	**(Acetaminophen as basis)**						
Ranitidine (Acid reducer)	3.208	1.238	2.592	2121	0.048	0.017	6.398
Phenylephrine (Nasal decongestant)	0.969	1.124	0.862	2071	0.778	-1.553	3.491
Omeprazole (Acid reducer)	2.320	1.182	1.963	2112	0.199	-0.634	5.275
Naproxen (Pain reliever)	0.717	1.110	0.646	2065	0.778	-1.647	3.081
Ibuprofen (Pain reliever)	6.567	1.416	4.638	1779	2.64E-05	2.754	10.381
Guaifenesin (Cough suppressant)	-4.029	1.003	-4.017	1961	0.000	-6.678	-1.380
Cimetidine (Acid reducer)	-1.497	1.064	-1.408	2092	0.478	-4.046	1.051

The sequential Bonferroni adjusted significance level is .05. Confidence interval bounds are approximate.

For the fixed effects of question type, participants response was significantly slower when answering the questions with "yes" as correct answer (ME=21.476, SE=1.551) with a difference of 7.528 seconds (SE=0.789, $p<0.001$) than the questions with "no" as correct answer (ME=13.947, SE=1.004). This, too, had mediating factors that were present in the form of significant interaction terms.

To examine the effect of the ingredient being tested on the response time, contrasts compared trials of a particular ingredient to trial times from acetaminophen (ME=16.528,

SE=1.375). Results suggested that accurate responses were made significantly quicker 4.029 seconds (SE=1.003, p<0.001) for the trials with guaifenesin (ME=12.499, SE=1.02); but significantly slower 6.567 seconds (SE=1.416, p=2.64E-05) for the trials involving ibuprofen (ME=23.095, SE=1.921), ranitidine trials were also significantly slower than the base by 3.208 seconds (SE=1.238, p=0.048) (ME=19.736, SE=1.639). No significant differences in response time were found in any other contrast comparisons involving ingredients.

4.5.3.2 Significant 2-way Interaction: PerLab x Layout

To test the hypothesis of this study, to analyze the interaction between concept educated and control groups (PerLab condition) and layout is of importance. Pairwise comparisons were conducted to interpret this interaction and results are shown in Figure 4.7 and Table 4.13-4.15. When the personalized labeling concept was introduced prior to the test (Table 4.14), participants were significantly faster making correct decisions than the control group (9.467 seconds (p=4.44E-16<0.001)) when trials involving personalized FOP layouts were compared (ME=9.584, SE=0.854) than the trials with the standard layout (ME=19.052, SE=1.707). In contrast, for the control group (Table 4.14), there was no evidence of a significant difference of response time to correct answers between the trials with personalized FOP layouts (ME=22.1, SE=2.052) were compared to standard formats (ME=22.234, SE=2.064) (p=0.897). From another perspective, when the personalized FOP was depicted (Table 4.15), people who were educated with the personalized labeling concept (ME=9.584, SE=0.854) used 12.515 seconds lesser (SE=2.013, p=2.44E-08<0.001) significantly for deciding correct answers than participants in the control group (ME=22.1, SE=2.052), who had not been acquainted to the concept prior to the study. In contrast, for the standard commercial layouts, no evidence of significant difference was present was found when the response accuracy of those introduced to the concept of

personalized labeling (ME=19.052, SE=2.322) was compared to those in the control group (ME=22.234, SE=2.064). (p=0.175)

Table 4.13 Results of Estimated Means of the Interaction between PerLab and Layout

PerLab	*Layout*	*Mean*	*Std. Error*	*95% Confidence Interval*	
				Lower	Upper
Concept-educated Group	Personalized Front-of-Pack	9.584	0.854	8.025	11.446
	Standard	19.052	1.707	15.938	22.773
Control Group	Personalized Front-of-Pack	22.1	2.052	18.367	26.59
	Standard	22.234	2.064	18.48	26.751

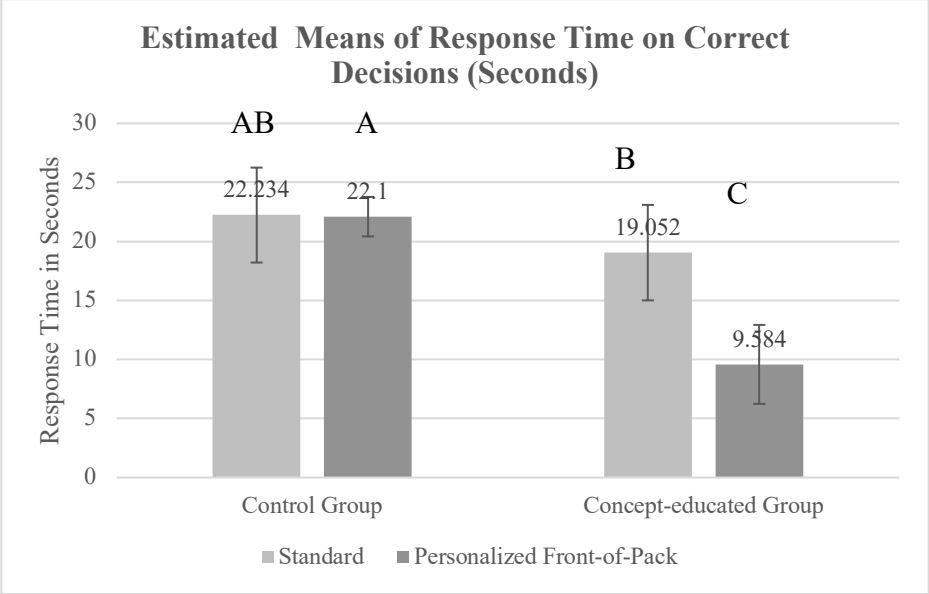


Figure 4.7 Estimated Means of Response Time on Correct Decisions

Table 4.14 Pairwise Comparisons of Effect of Layout on Response Time under each PerLab level.

PerLab	*Layout Pairwise Contrasts*	*Contrast Estimate*	*Std. Error*	*t*	*df*	*Adj. Sig.*	*95% Confidence Interval*	
							Lower	Upper
Concept-educated Group	Personalized Front-of-Pack vs. Standard	-9.467	1.057	-8.957	179	4.44E-16	-11.553	-7.382
Control Group	Personalized Front-of-Pack vs. Standard	-0.134	1.035	-0.13	2057	0.897	-2.164	1.896

The sequential Bonferroni adjusted significance level is .05; Confidence interval bounds are approximate.

Table 4.15 Pairwise Comparisons of Effect of PerLab on Response Time under each Layout level.

Layout	*PerLab Pairwise Contrasts*	*Contrast Estimate*	*Std. Error*	*t*	*df*	*Adj. Sig.*	*95% Confidence Interval*	
							Lower	Upper
Personalized Front-of-Pack	Concept-educated Group vs. Control Group	-12.515	2.013	-6.217	77	2.44E-08	-16.524	-8.507
Standard	Concept-educated Group vs. Control Group	-3.182	2.322	-1.37	78	0.175	-7.805	1.441

The sequential Bonferroni adjusted significance level is .05; Confidence interval bounds are approximate.

4.5.3.3 Significant 3-way Interaction: PerLab x Layout x Question Type

To interpret the three-way interaction that resulted in a significant effect on time to correct response which occurred between the concept educated (Perlab) x design layout (personalized FOP vs standard) x question type (affirmative or negative response to appropriateness for a given scenario), pairwise comparisons were analyzed. Figure 4.8 and Table 4.16-4.19 present findings.

Table 4.16 Results of Estimated Means of Response Time on Correct Decision Making

PerLab	*Layout*	*QuestionType*	*Mean*	*Std. Error*	*95% Confidence Interval*	
					Lower	Upper
Concept-educated Group	Personalized Front-of-Pack	Question with "Yes" as correct answer	10.333	0.979	8.561	12.470
		Question with "No" as correct answer	8.890	0.842	7.365	10.730
	Standard	Question with "Yes" as correct answer	24.979	2.388	20.661	30.198
		Question with "No" as correct answer	14.531	1.387	12.023	17.562
Control Group	Personalized Front-of-Pack	Question with "Yes" as correct answer	28.115	2.778	23.107	34.207
		Question with "No" as correct answer	17.371	1.707	14.291	21.115
	Standard	Question with "Yes" as correct answer	29.313	2.895	24.095	35.661
		Question with "No" as correct answer	16.864	1.658	13.873	20.500

Covariates appearing in the model are fixed at the following values: Age=34.42

For all comparisons related the question type (appropriate yes and no) across the other variables- the concept educated group + personalized FOP; the concept-educated group + standard; control group + personalized FOP; control group + standard), participants took significantly longer to correctly respond to the affirmative questions- “Yes, appropriate” compared to the negative “not appropriate” as the correct response. This is consistent with the findings of others, which supports the notion that target absent searches take significantly longer

than target present. The detailed results of pairwise comparisons were shown in the Figure 4.8 and Table 4.17.

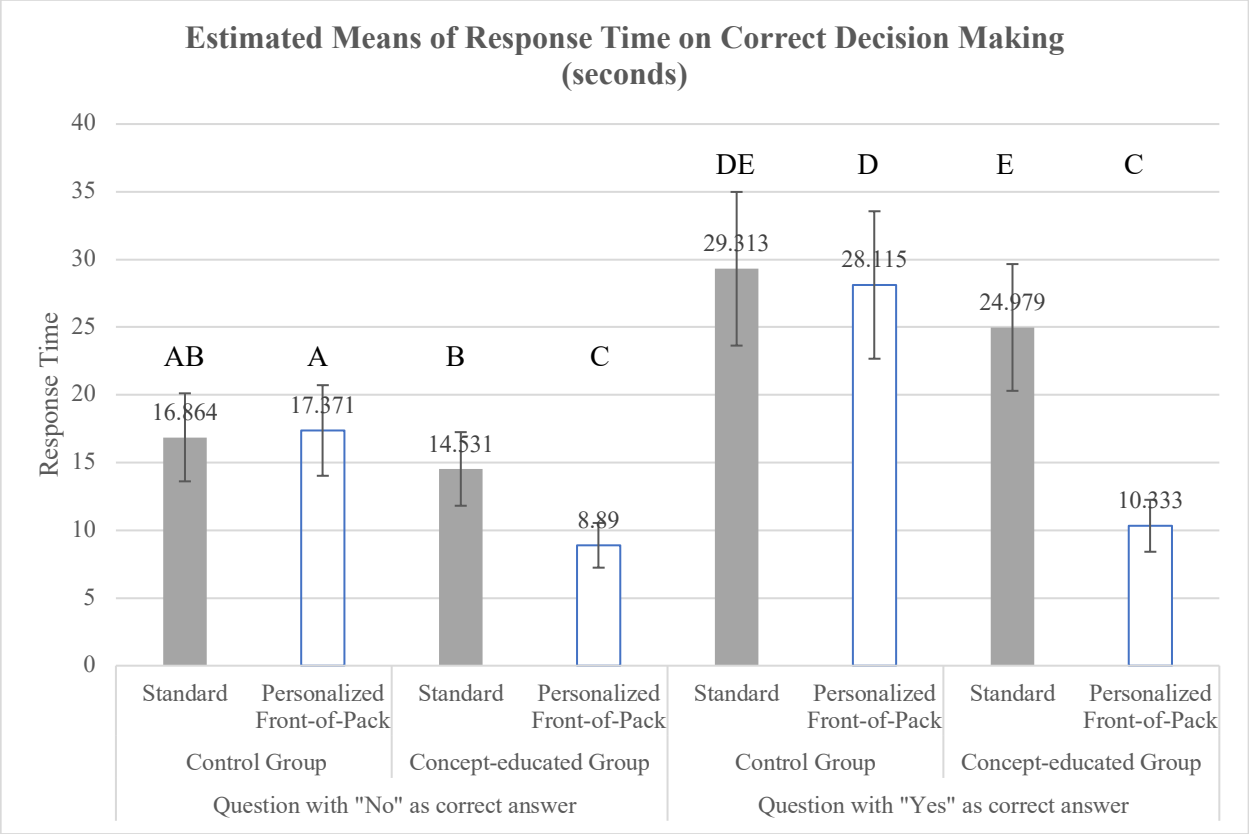


Figure 4.8 Estimated Means of Response Time on Correct Decision Making (Seconds)

Data also supported the idea that people in the group that were concept-educated took significantly less time to make decisions than people in the control group regardless of whether the correct question response was affirmative (appropriate for the scenario) or negative (not appropriate for the given scenario). Details relating to the pairwise comparisons for this analysis are shown in the Figure 4.8 and Table 4.18. The benefit of the personalized labeling strategy was most notable, when answering questions with “Yes” after having been educated to the concept (ME=10.333, SE=0.979). This group made decisions 17.782 seconds faster, on average, than people in the control group (ME=28.115, SE=2.778).

Furthermore, we also compared the effect of personalized FOP with the effect of standard labeling layouts within each level of combination of Question Type and concept educated (Perlab). Results did not indicate evidence of significant difference in response time to correct response when trials with personalized FOP were compared with trials employing the standard label were compared within the control group, the group not made aware of the concept ($p=0.530$; ME=28.115, SE=2.778 vs ME=29.313, SE=2.895 respectively). This was not the case for the concept educated group, where a significant difference for this comparison was noted. For the concept-educated group, more specifically, when trial questions were "yes" as correct answer, people spent 14.646 seconds lesser in average ($p=6.66E-15<0.001$) when the presence of personalized FOP layout (ME=10.333, SE=0.979) than the presence of standard layout (ME=24.979, SE=2.388); similarly, when trial questions were "no" as correct answer, people made decisions 5.641 seconds faster in average ($p=1.53E-09<0.001$) when the presence of personalized FOP (ME=8.890, SE=0.842) than the presence of standard layout (ME=14.531, SE=1.387).

Table 4.17 Pairwise Comparisons of Question on a certain level of PerLab and Layout

PerLab	Layout	Question Type		Mean Difference	Std. Error	t	df	Adj. Sig.	95% Confidence Interval	
		(Column A)	(Column B)	(A-B)					Lower	Upper
Concept-educated Group	Personalized Front-of-Pack	Question with "Yes" as correct answer	Question with "No" as correct answer	1.443	0.632	2.284	2116	0.022	0.204	2.681
	Standard	Question with "Yes" as correct answer	Question with "No" as correct answer	10.448	1.613	6.475	575	2.03E-10	7.279	13.617
Control Group	Personalized Front-of-Pack	Question with "Yes" as correct answer	Question with "No" as correct answer	10.744	1.81	5.935	673	4.70E-09	7.189	14.298
	Standard	Question with "Yes" as correct answer	Question with "No" as correct answer	12.449	1.921	6.481	507	2.16E-10	8.676	16.223

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable response time.
The mean difference is significant at the .05 level.

Table 4.18 Pairwise Comparisons of PerLab on a certain level of Layout and QuestionType

Layout	Question Type	PerLab		Mean Difference	Std. Error	t	df	Adj. Sig.	95% Confidence Interval	
		(Column A)	(Column B)	(A-B)					Lower	Upper
Personalized Front-of-Pack	Question with "Yes" as correct answer	Concept-educated Group	Control Group	-17.782	2.73	-6.514	102	2.81E-09	-23.197	-12.367
	Question with "No" as correct answer	Concept-educated Group	Control Group	-8.481	1.726	-4.913	102	3.42E-06	-11.906	-5.057
Standard	Question with "Yes" as correct answer	Concept-educated Group	Control Group	-4.335	3.316	-1.307	108	0.194	-10.908	2.238
	Question with "No" as correct answer	Concept-educated Group	Control Group	-2.333	1.91	-1.222	107	0.225	-6.119	1.453

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable Response Time
a The mean difference is significant at the .05 level.

Table 4.19 Pairwise Comparisons of Layout on a certain level of PerLab and Question Type

PerLab	Question Type	Layout		Mean Difference	Std. Error	t	df	Sig.	95% Confidence Interval	
		(Column A)	(Column B)	(A-B)					Lower	Upper
Concept-educated Group	Question with "Yes" as correct answer	Personalized Front-of-Pack	Standard	-14.646	1.758	-8.331	235	6.66E-15	-18.11	-11.183
	Question with "No" as correct answer	Personalized Front-of-Pack	Standard	-5.641	0.921	-6.124	685	1.53E-09	-7.449	-3.833
Control Group	Question with "Yes" as correct answer	Personalized Front-of-Pack	Standard	-1.199	1.909	-0.628	2069	0.53	-4.943	2.546
	Question with "No" as correct answer	Personalized Front-of-Pack	Standard	-0.507	1.125	-0.451	2063	0.652	-2.714	1.699

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable ResponseTime
a The mean difference is significant at the .05 level.

4.6 Discussion of this Chapter

For this study, our research goal was to develop pilot data regarding a theorized, personalized FOP labeling strategy. To test the strategy, we employed objective measures intended to serve as proxies for the effectiveness (accuracy) and efficiency (time to make decisions) related to people's decisions when evaluating an OTC given a specific scenario.

Results of significant 2-way interactions for both response accuracy (Figure 4.6 and Chapter 4.5.2.2) and response time to correct answers (Figure 4.8 and Chapter 4.5.2.3) suggest that consistent with our hypothesis, when people are educated about the novel labeling concept, the personalized FOP labeling strategy is both effective (the probability of providing correct answers) and efficient (the time to spent on making correct decisions) on helping people make decisions given a specific scenario as compared to the standard labels.

The effectiveness of the novel approach was characterized by the response accuracy of participants. The personalized approach improved both response accuracy compared to the standard label layout (Figure 4.6) as well as the time to response (Figure 4.8) among those that were aware of the approach. By contrast, the benefit of the personalized strategy yielded no benefit to accuracy over the standard labeling layout among those that were not informed of the approach. Pairwise comparisons of accuracy across the groups educated related to the concept by label design suggested that the only significant difference in accuracy for all comparisons happened for those educated about the strategy and viewing personalized labels (See numerical results in Chapter 4.5.2.2). That is, the standard label performed equivalently in the concept educated group and those not educated to the concept, with the personalized concept showing a similar performance in accuracy when the concept was not explained to participants. This

shouldn't be a problem for our proposed approach as only people that were aware of the concept would be able to install an application to their device.

We utilized the dependent variable "time to a correct assessment" as a proxy for efficiency. As with accuracy, whether the participant had been informed of the new labeling strategy, impacted its effect on performance as measured by time to correct selection. (Chapter 4.5.2.2. Consistent with our hypothesis, people that were informed about the personalized labeling strategy made decisions significantly faster than the control group for the trials with the presence of the personalized FOP. And, as expected, for the trials with standard commercial layouts, no significant differences were identified. (Figure 4.6) (See numerical results in Chapter 4.5.2.2)

A post-hoc analysis of the three-way interaction further investigated the mediating effect of question type (whether an affirmative or negative response was correct) on the response time of correct decisions. (See numerical results in Chapter 4.5.2.3) Evidence suggested that people took less time to find answers when "no" was the correct response as compared to those trials where "yes" as correct. (Table 4.11) This finding is supported by the literature, which suggests that searching for an absent target (there is no warning present suggesting not to take this product) generally result in longer search times than target present searches. (Goldman, 2018; Robin, 2015). The finding is also intuitive, as the questions were designed such that those with "no" as correct answers directly corresponded to key words from the box of the DFLs on the packages. Using Tylenol for example, we asked, "Is this pain reliever appropriate to uses for a person with a fatty liver?", where the key word "fatty liver" corresponded to the "liver warning" and "ask a doctor before use if liver disease" on the DFL. Conversely, questions with an affirmative response, would not contain directly correlated information in the DFL. For the

Tylenol example again, we asked “is this pain reliever appropriate to use for a person taking melatonin?” In this case, the key words “taking melatonin” cannot be directly found corresponding information on the warning label. Participants needed to eliminate the possibility that melatonin was contraindicated with acetaminophen. Notably, in the real-world situation, the decision-making process could be similar to the trials of questions with “yes” as correct answers, where the key words cannot be found correspondingly on the DFL of packages, or even the key words themselves were rarely formed in patients’ mind due to their limitations of professionalism in the medicine domain.

Additionally, evidence suggested that the effects of ingredients also have significant impacts on both effectiveness (accuracy, see Chapter 4.5.2.3) and efficiency (time to make decisions, see Chapter 4.5.3.1) related to people’s decisions. It is somewhat academic in nature, since a change of an active ingredient involves complete changes of commercial brand, packaging design, drug facts label as well as trial question designs. Participants’ performance on response accuracy and decision time could be possibly impacted by their familiarity of such active ingredients and brands, the color or layout of such packaging designs, the text length of such drug facts labels as well as the difficulty of the related trial questions.

Chapter 5 Discussion

5.1 Summary and Interpretations

Our broad research goal was to provide information that could be used to improve interactions between self-medicating consumers. We theorized that a novel labeling strategy leveraging emerging technologies would be helpful in improving all stages of information processing for consumers considering an OTC product for their own use. Labeling is an important means of communicating critical information to consumers as it is generally available at the point of purchase and consumption and widely accessible. As such, the quality of interactions between the label of an OTC and consumers play an important role for safe and effective use of medications. Under the framework of Human-Package Interaction model, our studies firstly investigated the efficacy of the non-directive FOP OTC labeling strategy on consumer's attentive behaviors, and progressively proposed the concept of directive, personalized FOP labeling strategy. We develop pilot data which supported the efficiency and efficacy of this approach.

5.1.1 Non-directive FOP Labeling Strategy on Attention Garnering

In our change detection testing of the non-directive FOP OTC labeling strategy, our results suggested that including critical warnings in an FOP box increased both the effectiveness and efficiency related to attention to critical information. By presenting the critical information inside of FOP box on PDPs, the probabilities of such information to be processed by consumers was enhanced. This is promising because attention is a prerequisite of further processing under the H-PIM paradigm.

The efficacy of highlighting content critical to the consumer's safe and effective use of these products was clear; having benefitted both efficiency (speed) and effectiveness (accuracy).

Highlighting attracts consumers' attention to the content and hence increases the possibility for such content to be noticed as well as processed. The conclusion finding aligns well with previous studies focused on the benefits of highlighting a specific active ingredient (King et al., 2011) and recent changes to US regulations which require the same.

5.1.2 Personalized FOP Labeling Strategy on Decision Making

In the absolute judgement test of the personalized FOP labeling strategy, the results suggested that consumers who were educated regarding the personalized labeling strategy would make decisions more appropriately as well as spend less time making appropriate decisions for trials comprised of the novel label system compared the ones which employed the existing, commercial approach. In short, the personalized FOP labeling strategy shows the promise of being more effective (ability to make the correct decision) and efficient (time to correct decision) during late-stage information processing (decision making) if consumers are aware of the strategy. Given the proposed approach (an application to be installed on the consumer's phone) the issue of concept education is moot.

5.2 Implications

5.2.1 Discussion in Context of Theories

Related to the Human Package Interaction model discussed in Chapter 1.5, the studies in this dissertation were primarily focused on early-stage processing (perception) and late-stage processing (comprehension). Our change detection strategy flattened out the stimuli, providing a conservative measure which biased against our design by "forcing" the DFL to be exposed.

For the early stages of information processing, a change detection test was utilized to examine the consumers' allocation of attention to critical information required for the safe and effective use of OTC labels. The allocation of attention indicates how consumers' focus on the

OTC label design elements, and hence affects the likelihood of their perception of such information in the bottom-up mechanism. It is interesting to note that our attentional system not only allows us focus on something considered as important, but also "tune out" information that are not relevant at that moment. In one way, when the information was directly highlighted, more attention was allocated on the highlighted changes, comparing to those were not highlighted. In another way, when the front-of-pack label was presented, the effect of attention garnering of such labels drove consumers "tune out" the information, which was blinking, outside of such label.

For the later stages of information processing, the absolute judgement task investigated the efficacy of personalized FOP labeling strategy on participants' decision making (late-stage processing). The study found the personalized FOP labeling strategy to be useful for facilitating later stage processing action, specifically to make decisions of product appropriateness under specific scenarios related to the critical active ingredient or warning information.

5.2.2 Discussion in Context of Existing Knowledges

For the change detection test of non-directive FOP labeling strategy, there are two previous studies close to the experimental design of this study. Esfahanian et al. (2020) conducted an FOP labeling change detection study using mock brands of OTCs with participants 65 years and older. Harben et al. (2021) also focused on examining the efficacy of FOP labeling strategy on the efficacy of FOP labeling strategy (but with bottom right corner on the PDP) among senior adults.

Esfahanian's study affirmed highlighting as a promising strategy for improving the detection accuracy as well as shortening the response time for detecting changes successfully. Also, no evidence was found in the study to conclude that the presence of an FOP, compared to a

standard label, statistically affected attention garnering or time to detect critical information for OTC products among adults with 65 years old or older. In contrast, our results in the change detection test suggested that the presence of an FOP prolonged the detection time than the standard ones, under the condition that the changes were outside of the FOP box, but the information (DD1+DD2) changed inside of the FOP box attracted more attention. To interpret this difference, it is worth noting that the major difference between two studies is the age of testing population, where Esfahanian focused on older adults (65+), our study recruited from young adults in their 20s and 30s. It is possible that older adults were more familiar with the layouts of medical packaging, such as DFL, due to an enhanced propensity to engage in polypharmacy as compared to their younger counterparts. And hence older adults had already established their way to read such packages label with less attention on the bottom-left corner of PDP.

Harben's study further supported the use of highlighting as a helpful method in attracting older adult's attention to information. Additionally, the study also concluded that the drug interaction warnings in the FOP label (on the package's front panel) attracted more attention than that on the DFL regarding the time to detect changes. Those results are in line with our study. However, it is worth noting that there were several differences in the experimental designs and testing population between two studies. Firstly, the non-directive FOP label as well as the PDP designs in Harben's study were different. Comparing to our design, the non-directive FOP label in Harben's study used a larger box both in height and width. And the line space was doubled between different drug interaction warning information inside of the FOP label box, while ours were single spaced with bullet points. Additionally, and most importantly, the location of the FOP labels was on the bottom-right corner of the PDP in Harben's study, where the FOP was

nearly in the center of the stilmui and closer to the DFL. And the active ingredient information on the PDP was, in exchange, replaced to the left side on the PDP. Also, for the testing population, Harben's study focused on older adults aged 65 years old or older. It is worth noting that, even though with the above differences between the two studies, the results of the two studies were consistent, suggesting this to be a promising design strategy for enhancing the attention to, and use of, critical information.

5.3 Limitations

Despite the meaningful results found in this dissertation, there are limitations associated with any study, and this one is no different.

The generalizability could be limited by the stimuli used in the study. Specifically, three OTC mock brands from three drug categories were used in the change detection test, while nine real-world OTC brands from the same three drug categories were employed in the absolute judgement tests. Even though we carefully chose the drug ingredients, brands and drug categories to be representative products for the OTC markets, we must acknowledge that the results from only three categories limited the generalizability to the broader OTC market.

Additionally, for the mock brands in the change detection test, grayscale was utilized in designing those labels. This was done as pilot work which controlled for the effect of color on the attention allocation. That said, this (obviously) does not represent realistic conditions and undoubtedly enhanced the visual salience of the highlighted treatments. Moreover, for both studies, we only tested the labels in the flatten format, but in the real-world situation the packaging containing this information would be physical forms where consumers would have to provide some type of action to access information present in the DFL (turning a box or a bottle).

That said, we believe our approach biases toward the accessing the DFL, away from benefiting our proposed solution, yielding a conservative estimate of the benefit of our proposed design.

Research design was another limitation in our research. For both studies in this dissertation, the lack of adding a post-test questionnaire which collected information about prior familiarity related to drug ingredients limits our ability to single out the effect of drug ingredient familiarity when investigating the effects of FOP labels on attention garnering and decision making.

Moreover, for the treatments that included the personalized labeling strategy during the absolute judgement test, we utilized a simple, computer-based task, rather than employing an augmented app via smartphone. As such, it could be argued that results present an absolute maximum benefit that could be obtained if such an approach were employed, as the technology could prove cumbersome or difficult for participants and would require that they disclose (and update) full and accurate histories of their conditions and medications.

5.5 Conclusions

In conclusion, the work from this dissertation theoretically proposed directions for developing labeling strategies to improve interactions between consumers and OTC products based on the framework of human-package interaction model, namely non-directive, semi-directive and personalized (directive) FOP labeling strategies. This dissertation also involved investigations of the effect of non-directive FOP labeling format on attention garnering in the early stages of information processing, as well as the effect of personalized FOP labeling strategy on decision support for consumers to make appropriate choices. Those work support further investigations of the FOP labeling strategies for both non-directive and personalized (directive).

5.6 Suggestions for the Future Study

We are in the era of metaverse and the surge of artificial intelligence technologies. The fast-paced iterations of augmented reality technologies and smart devices could ignite a new direction of OTC labeling strategies: a smart, personalized, and fast reacting OTC augmented FOP labeling system. This system could serve a very important role in improving the consumer-package interactions in the OTC market, because it could assist consumers to make decisions appropriately and quickly based on their own health conditions when healthcare professionals are not available.

The results from our preliminary investigations on the personalized FOP labeling strategy suggested its efficacy and efficiency on consumers' decision making for the avatars under given healthcare scenarios. However, to further improve the personalized FOP labeling systems, studies are needed to fill the gap of knowledge under the following directions.

Firstly, future studies can focus on the improvement and development of personalized labeling strategies. In this dissertation, we only examined the effect of personalized labeling on decision making in the later stages of information processing. For the other stages such as attention allocation and comprehension, future work is needed to examine the impacts of the following areas: (1) design elements of personalized labeling symbols, such as shape, size, color, and border thickness; (2) presenting methods of personalized labels, such as flashing, animating, warning sounds, and even voices or short videos from your doctors.

Secondly, to extend the generalization of personalized labeling strategies, researchers should further investigate the effects of such strategies in the real-world settings, involving: (1) to examine the effectiveness of personalized labeling through augmented reality via smart devices. It is necessary because the interactions between consumers and augmented personalized

labeling via smartphone app are very different from the interactions with the flattened mock personalized labeling concept in our studies. (2) to expand scope of research from folding cartons to other packaging types, such as bottles, pouches. It is because that the degrees of distortion of PDPs and DFLs on other packaging types rather than boxes could impact the exposure of information in the early stages of information processing. (3) to take more interactive scenes into considerations, such as in-store settings and at-home settings.

Thirdly, future studies should also consider the effectiveness of personalized labeling strategies on populations with different needs. The behaviors of older adults, children, pregnant women and disable people could be largely different than general population recruited in this study. It is necessary to make the personalized labeling designs accessible to all people, regardless of age, disability, or other factors. Moreover, as the education of the concept of the personalized labeling strategies is as important as the strategies themselves, the way to educate such concept should also be improved to make it more acceptable among broader populations effectively and efficiently.

APPENDICES

APPENDIX A Stimulus Designs for the Change Detection Test

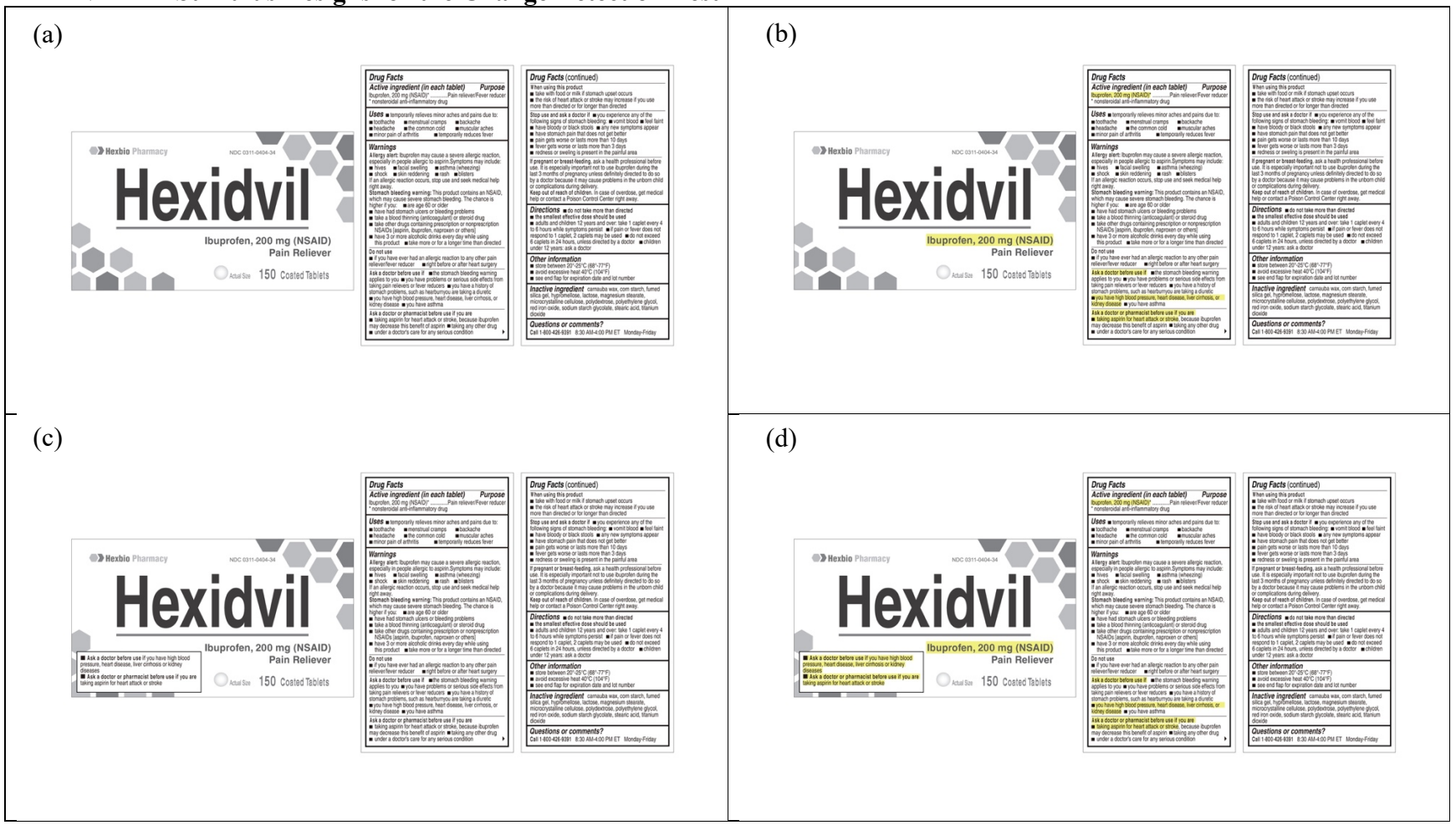


Figure X.1 A Group of Hexidvil (Ibuprofen) Treatments. (a) Standard, no Highlight; (b) Standard, Highlight; (c) FOP, no Highlight; (d) FOP, Highlight.

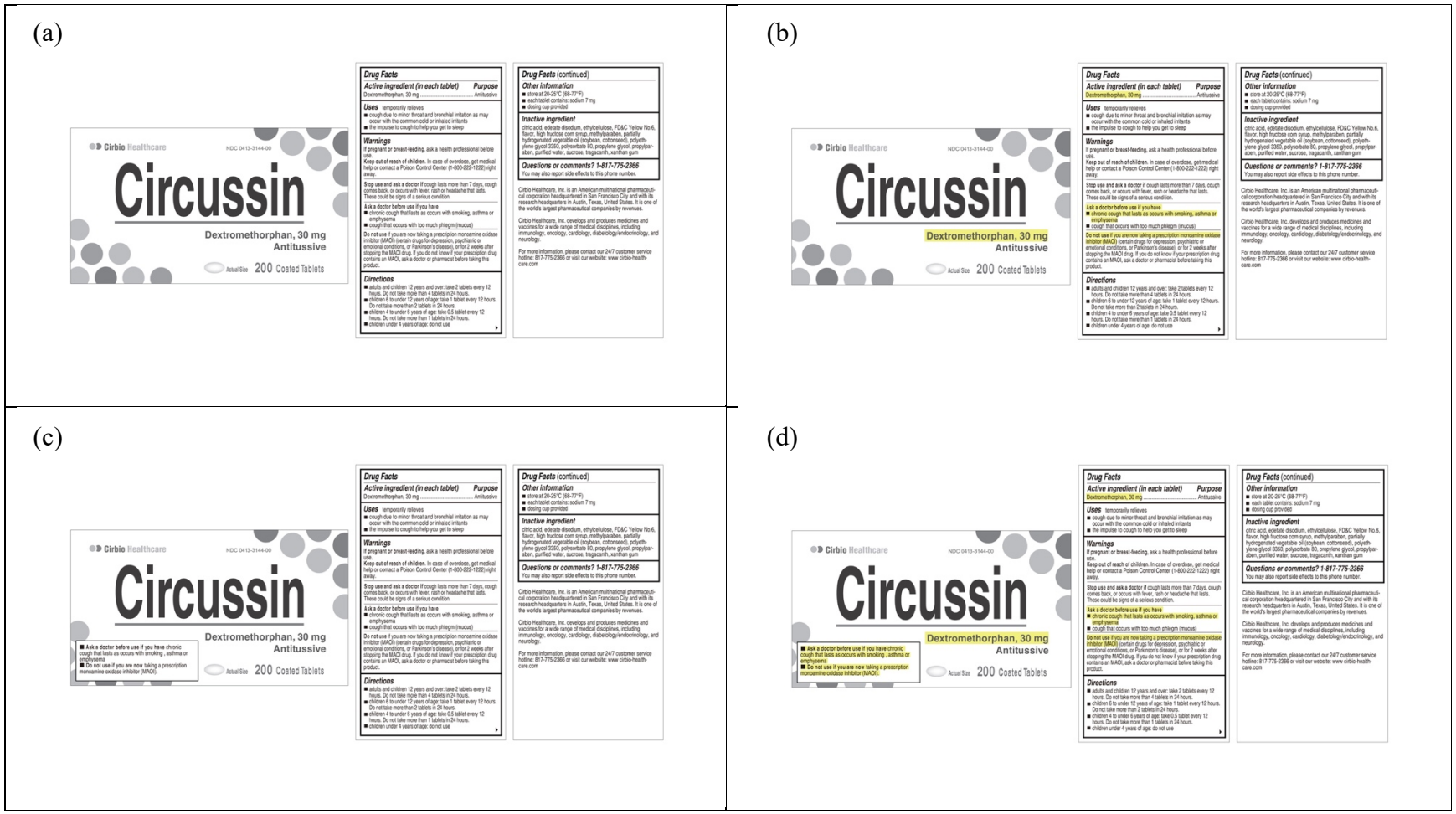


Figure X.2 A Group of Circussin (Dextromethorphan) Treatments. (a) Standard, no Highlight; (b) Standard, Highlight; (c) FOP, no Highlight; (d) FOP, Highlight.

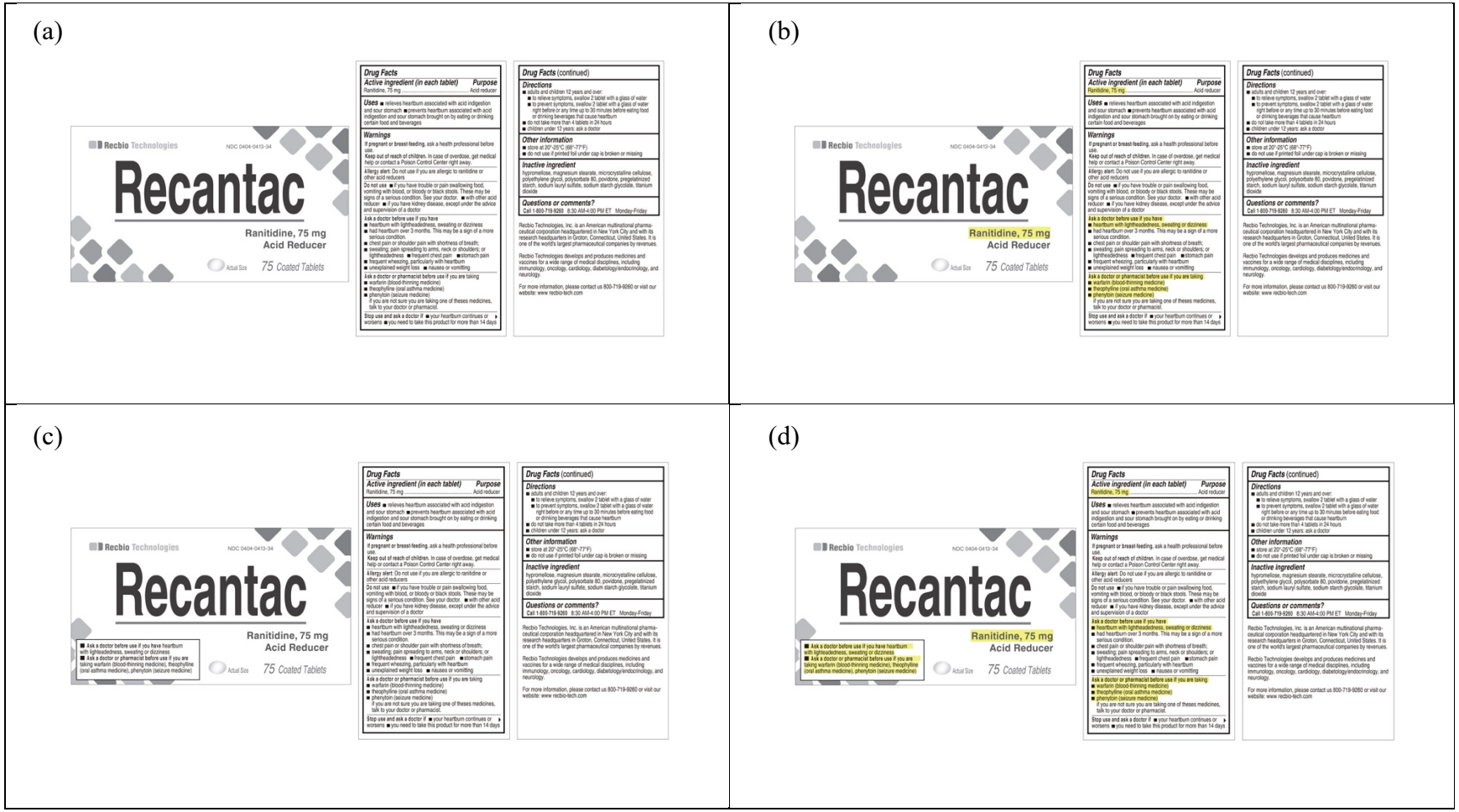


Figure X.3 A Group of Recantac (Ranitidine) Treatments. (a) Standard, no Highlight; (b) Standard, Highlight; (c) FOP, no Highlight; (d) FOP, Highlight.

APPENDIX B IRB approved flyer for change detection recruitment.

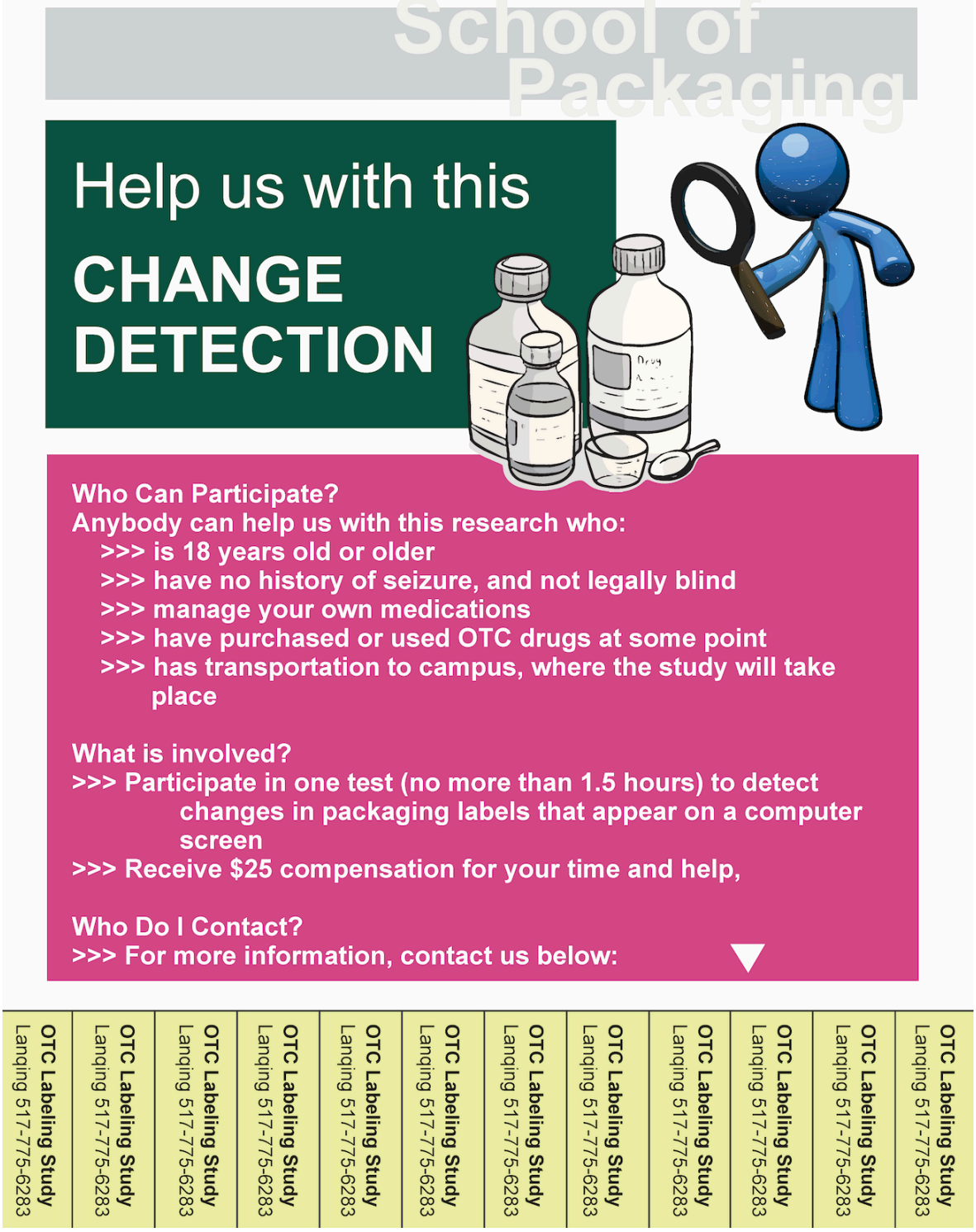


Figure X.4 The Recruitment Flyer used for Change Detection Test

APPENDIX C IRB approved consent form for FOP labeling change detection test

Michigan State University
School of Packaging/Department of Psychology
Study Title: Experiment 1- Change Detection – OTC Labels

INSTRUCTIONS AND RESEARCH CONSENT FORM

You are being asked to participate in a research project. Researchers are required to provide a consent form to inform you about the study, to convey that participation is voluntary, to explain risks and benefits of participation, and to empower you to make an informed decision. You should feel free to ask the researchers any questions or concerns you may have during the experiment.

1. **PURPOSE OF RESEARCH:** You are being asked to participate in an experiment to investigate how well the formats of different OTC drug labels work.

2. **TO PARTICIPATE IN THIS STUDY YOU MUST:**
   a. Be 18 years old or older
   b. Have NO HISTORY OF SEIZURE
   c. Not legally blind
   d. Manage your own medications
   e. Have purchased or used OTC drugs at some point
   f. Have transportation to Michigan State University, where testing will occur

2. WHAT YOU WILL DO: If you agree to participate in this study, the following events will take place. We will ask you to answer some basic questions about yourself. Your visual acuity, and your ability to see color will be tested; we will also test you for familiarity with medical terms. You will sit in front of a computer screen. On the computer screen, a test image (a label) continuously alternates with the same image, slightly altered, with a gray screen in between. This image-blank-test-blank will loop, resulting in a "flickering" at the place of alteration until you press the space bar, indicating that you have found the change. You will then be asked to use a mouse to point and click on the place where you saw the flickering of the image. If you are unfamiliar, unable or uncomfortable with using the mouse to click, you can point to the location of the change and the research team will do this for you. If you cannot find the change within the time the allotted time (30 seconds per label), the software will move you to the next trial in the test. This process will repeat for a series of trials. The research should take no more than 1 hour of your time. In exchange for your participation in this study, you will receive $25.

3. POTENTIAL BENEFITS: There will be no direct benefit to you from these procedures. However, it is our goal to understand what factors make certain parts of a label more noticeable than others so that we can develop labels that provide important information to people in ways that they are likely to see it.

3. **POTENTIAL RISKS:** We will ask you to read aloud a series of words used by medical people. It is possible that you may not be familiar with these words and this would be embarrassing. You can skip any words you are unsure of.

There is a possible risk of seizure associated with viewing flashing images; as a result, if you have a history of seizure, you are not eligible to participate. If you are injured as a result of your participation in this research project, researchers from Michigan State University will assist you in obtaining emergency care, if necessary, for your research-related injuries. If you have insurance for medical care, your insurance carrier will be billed in the ordinary manner. As with any medical insurance, any costs that are not covered or in excess of what are paid by your insurance, including deductibles, will be your responsibility.

The University's policy is not to provide financial compensation for lost wages, disability, pain or discomfort unless required by law to do so. This does not mean that you are giving up any legal rights you may have.

In the event that you are uncomfortable with any of the tasks, you may elect to skip a portion of the study, or discontinue altogether.

5. PRIVACY AND CONFIDENTIALITY: The data for this project will be tied to subject number, not name. Although the researchers, research staff, and the Institutional Review Board will have access to the data, neither the researchers nor anyone else will be able to link your data to you. Participant confidentiality will be protected to the maximum extent allowable by law. Paper records will be kept in Dr. Bix's office for as long as required by publishers or at least three years after the study closes whichever is longer; digital records will be housed on computers in our laboratories (Psychology and Packaging). Data would be provided (deidentified) to publications that deemed it a necessary part of due diligence and is also accessible to the IRB.

6. YOUR RIGHTS TO PARTICIPATE, SAY NO, OR WITHDRAW: Participation in this research is completely voluntary. Refusal to participate will involve no penalty or loss of benefit. You may also refuse to answer particular questions. You may change your mind at any time, for any reason, and withdraw without penalty or loss of compensation.

7. COSTS AND COMPENSATION FOR BEING IN THE STUDY: You will receive $25 in exchange for your participation in this study.

8. CONTACT INFORMATION FOR QUESTIONS AND CONCERNS: If you have any concerns or questions about this research study, such as scientific issues, how to do any part of it, or to report an injury, please contact the researcher (Laura Bix, PhD 448 Wilson, 114 Packaging East Lansing MI 48824 517 355-4556 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 408 W. Circle Drive, 207 Olds Hall, MSU, East Lansing, MI 48824.

9. DOCUMENTATION OF INFORMED CONSENT: I voluntarily agree to participate in the study.

_____ _____

Signature Date

You will be provided with a copy of this consent form for your records.

APPENDIX D Demographic questionnaire and Pretests.

Research Questionnaire Form Subject #: _____
Instructions for researcher to read are in **RED color** (but the color texts are shown in this ProQuest print as **"*Bold Italic*"** due to formatting requirements for colored text.)

Section A. Demographic Survey

1. **Sex: _____**

2. **Age:__**_____

3. **What is your ethnicity?**

□ White, non-Hispanic
□ Asian or Pacific Islanders
□ African Americans, non-Hispanic
□ American Indian/Alaskan Natives
□ Hispanic
□ Others: _____

4. **What is the highest level of education you have completed?**

□ Middle School
□ High School
□ Associate Degree
□ Bachelor's degree
□ Master's degree
□ Doctor Degree

5. **What is your native language?**

□ English □ Spanish □ French □ Russian □ Chinese
□ Japanese □ Others: _____

Section B. Near Point Visual Acuity and Health Literacy
Part I. Near Point Visual Acuity

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.

20/800: D T 4
20/400: L E S 3
20/250: R F X B N
20/200: P O 5 7 A
20/100: 8 C V L M
20/70: 3 7 S Z K
20/50: E X R T N
20/40: D M P R O F
20/30: F H G J X V
20/20: 3 A S R E P

Result: 20/____

Part II. REALM-R Examiner Record

"It would be helpful for us to get an idea of what medical words you are familiar with. What I need you to do is look at this list of words, beginning here [point to first word with pencil]***. Say all of the words you know. If you come to a word you don't know, you can sound it out or just skip it and go on."***

If the participant stops, say, ***"Look down this list*** **[point]** ***and say the other words you know."*** If the participant takes more than 5 seconds on a word, encourage the patient to move along by saying,

"Let's try the next word."

If the patient begins to miss every word or appears to be struggling or frustrated, tell the patient,

"Just look down the list and say the words you know."

Put an x next to the scored trials where subjects did not correctly pronounce the word and a checkmark next to those that were correctly indicated.

fat		fatigue	_____
flu		directed	_____
pill		colitis	_____
allergic	_____	constipation	_____
jaundice	_____	osteoporosis	_____
anemia	_____		

Fat, Flu, and Pill are not scored. We have previously used a score of 6 or less to identify patients at risk for poor literacy.

Score: _____

Section C. Color Differentiation Ability

Please hold each of these 75 cm (measure with string) ***from your eyes and read the number that appears to you. If no number is apparent, please say "pass".***

Write the number that the subject states for each trial on this form. Put an x through incorrect trials and a checkmark across the plates that are correct.

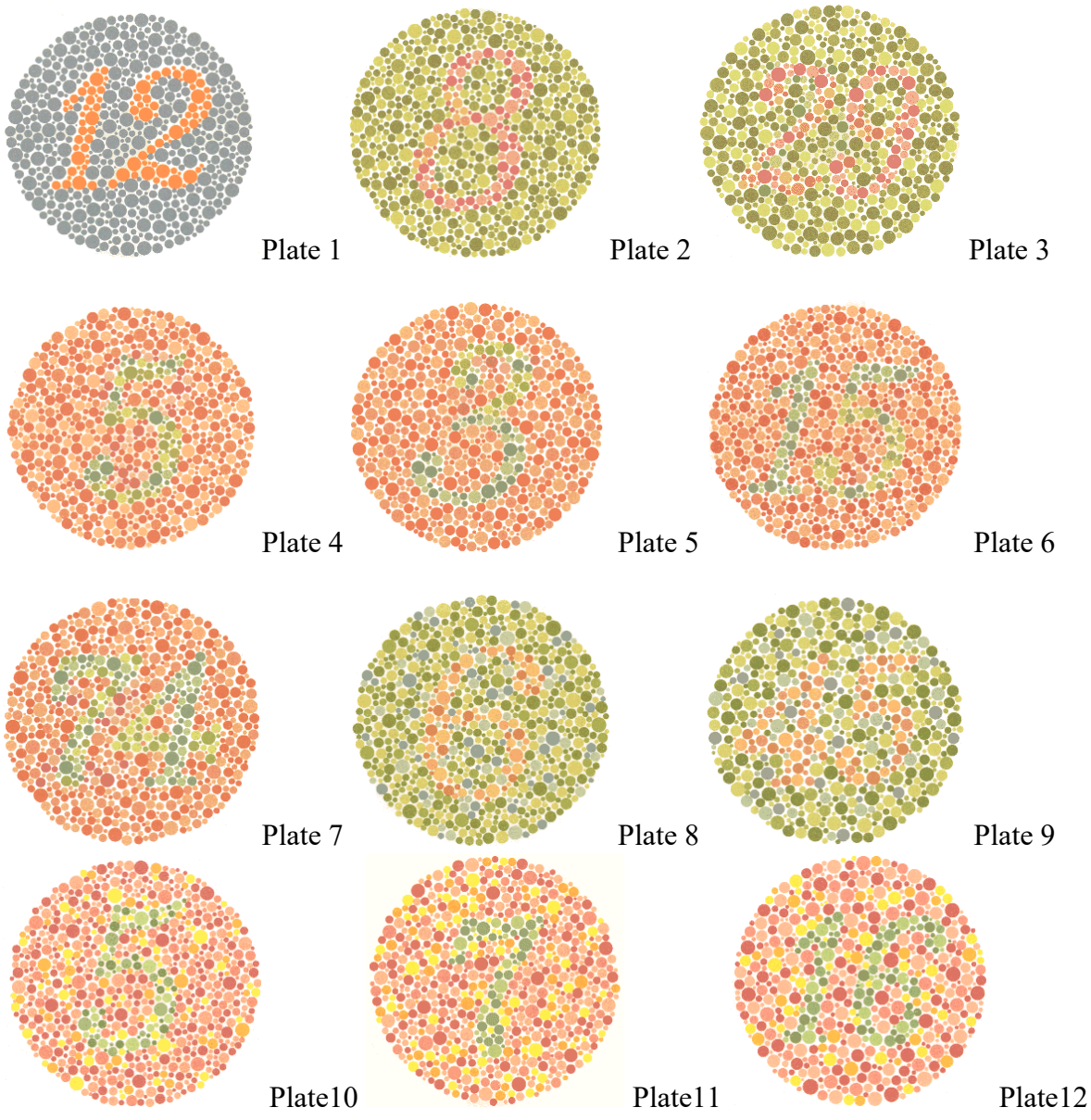


Figure X.5 The Plates for Color Differentiation Test

Figure X.5 (Cont'd)

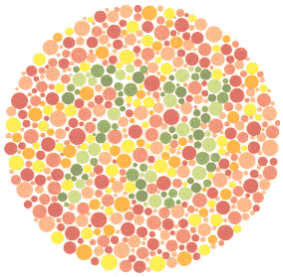


Plate13

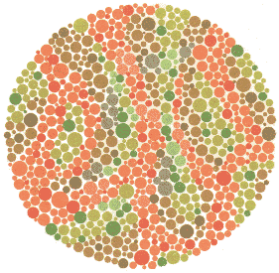


Plate14

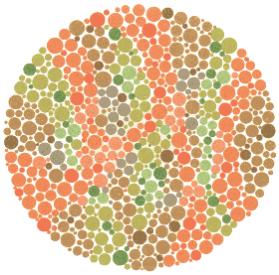


Plate15

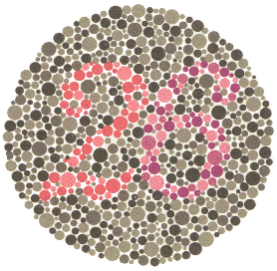


Plate16

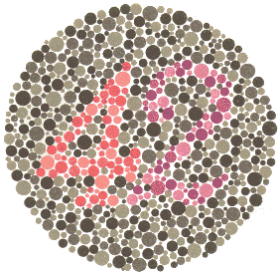


Plate17

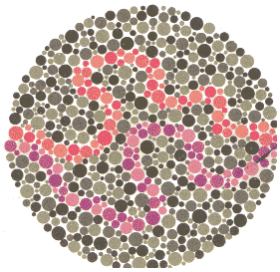


Plate18

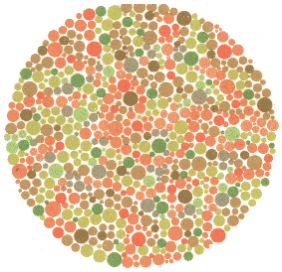


Plate19

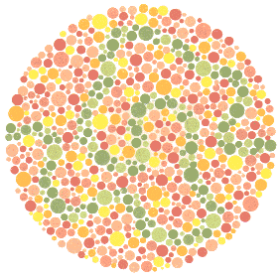


Plate20

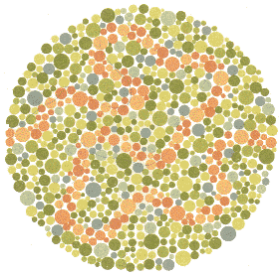


Plate21

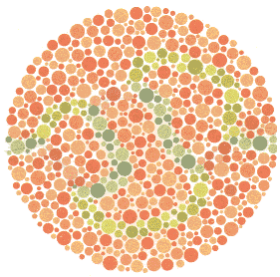


Plate22

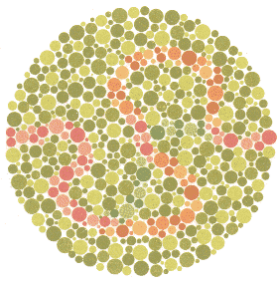


Plate23

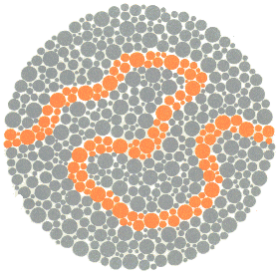


Plate24

Answers to each plate

Plate	Normal Person	Person with Red-Green Deficiencies				Person with Total Color Blindness and Weakness
1	12	12				12
2	8	3				X
3	29	70				X
4	5	2				X
5	3	5				X
6	15	17				X
7	74	21				X
8	6	X				X
9	45	X				X
10	5	X				X
11	7	X				X
12	16	X				X
13	73	X				X
14	X	5				X
15	X	45				X
		Protan		Deutan		
		Strong	Mild	Strong	Mild	
16	26	6	(2) 6	2	2 (6)	
17	42	2	(4) 2	4	4 (2)	

The mark X shows that the plate cannot be read. Blank spce denotes that the reading is indefinite. The numerals in parenthesis show that they can be read but they are comparatively unclear.

As assessment of the readings of plates 1 to 15 determines the normality or defectiveness of color vision.

If 13 or more plates are read normally, the color vision is regarded as normal.

If only 9 or less than 9 plates are read normally, the color vision is regarded as deficient.

However, in reference to plates 14 and 15, only those who read the numerals 5 and 45 and read them easier than those on plates 10 and 9 are recorded as abnormal readings.

It is rare to find a person whose recording of normal answers is 14-16 plates. An assessment of such a case requires the use of other color vision tests, including the anomaloscope.

APPENDIX E Details about Personalized Labeling Application

1. Personalized Labeling Application Github repository and source code

This program is available online via Lanqing Liu's Github repository: https://github.com/Lantrick-Liu/Personalized-Labeling-App. The repository contains a README.md file more detailed instructions to download, setup and use.

2. Xcode Documentations

a. UIKit Documentation: https://developer.apple.com/documentation/uikit

b. SceneKit Documentation: https://developer.apple.com/scenekit/

c. ARKit Documentation: https://developer.apple.com/documentation/arkit/

2. Four sample designs used in the application:

Figure X.6 The Flat Design of Advil Prototype

Figure X.7 The Flat Design of NyQuil Prototype

Figure X.8 The Flat Design of Tylenol Prototype

Figure X.9 The Flat Design of Pain Relief PM Prototype

APPENDIX F Instructions for Educating Personalized FOP Labeling Concept

Instructions for Participants with Personalized Labeling Education (New Version)

(After finishing the pretests, please place packages on the table, and setup the phone.)

Now we would like to introduce the concept of personalized labeling. Please open the app and point the camera to the packages.

As you can see, the personalized labeling concept augments (or add) a virtual symbol of risk level on the real-world packages. When the person is considering a product, the symbol indicates a decision about the appropriateness for use.

The symbol is changed based on the person's health history and other products that he or she is taking, so it gives a personalized suggestion which is just suitable for his or her situation.

As you can see, GREEN CHECK MARK indicates the drug is "appropriate-to-use" and RED CROSS OR STOP SIGN indicates "not-appropriate-to-use" for the person.

In the following test, a half of the trials you are about to undertake assume that the person in the scenario is using the app and it is returning a response that is specific to the drug and its appropriateness for the person. Half of the trials assume that the person is not using this app. Please answer the questions for the person in the trial as quickly as you can.

(Guide the participant to the laptop for the main test)

Instructions for Participants with Personalized Labeling Education (Old Version)

** Note that the data of participants who were educated with this version had been excluded from the final data analysis.

(After finishing the pretest) Now before we start the main test, we would like to introduce the concept of personalized labeling. We used augmented reality to build this concept and we would like to show you via a smartphone while the introduction.

So, firstly, let's look at some package prototypes on the table.

(After showing the packages on the table.) And now we have a smart phone here for you to use. And please click the app on the screen and use the camera to point at the package prototypes on the table. And let's see what have changed...

(Give participant the phone and open the app. After play with the phone, participants possibly say: Some marks on the upright package corner, check mark, stop sign mark and so on.) Yes, as you can see, the personalized labeling concept augments a layer of disk level on the real-world packages. The symbol can be changed based on user's health history to indicate the potential appropriateness of using the product. For example, it can be changed to check marks to indicate "appropriate-to-use" and stop signs for "not-appropriate". And in this way, the personalized labeling can potentially help on patient's medical decisions. So, this is the basic concept of personalized labeling.

Do you have any questions? (If yes, answer participants' questions… If not, then….)

Okay, let's go for the next step for the main test. (Guide the participant to the laptop for the main test)

APPENDIX G Trial Designs for Absolute Judgement Test

Figure X.10 The Practice Trials of the Absolute Judgement Test

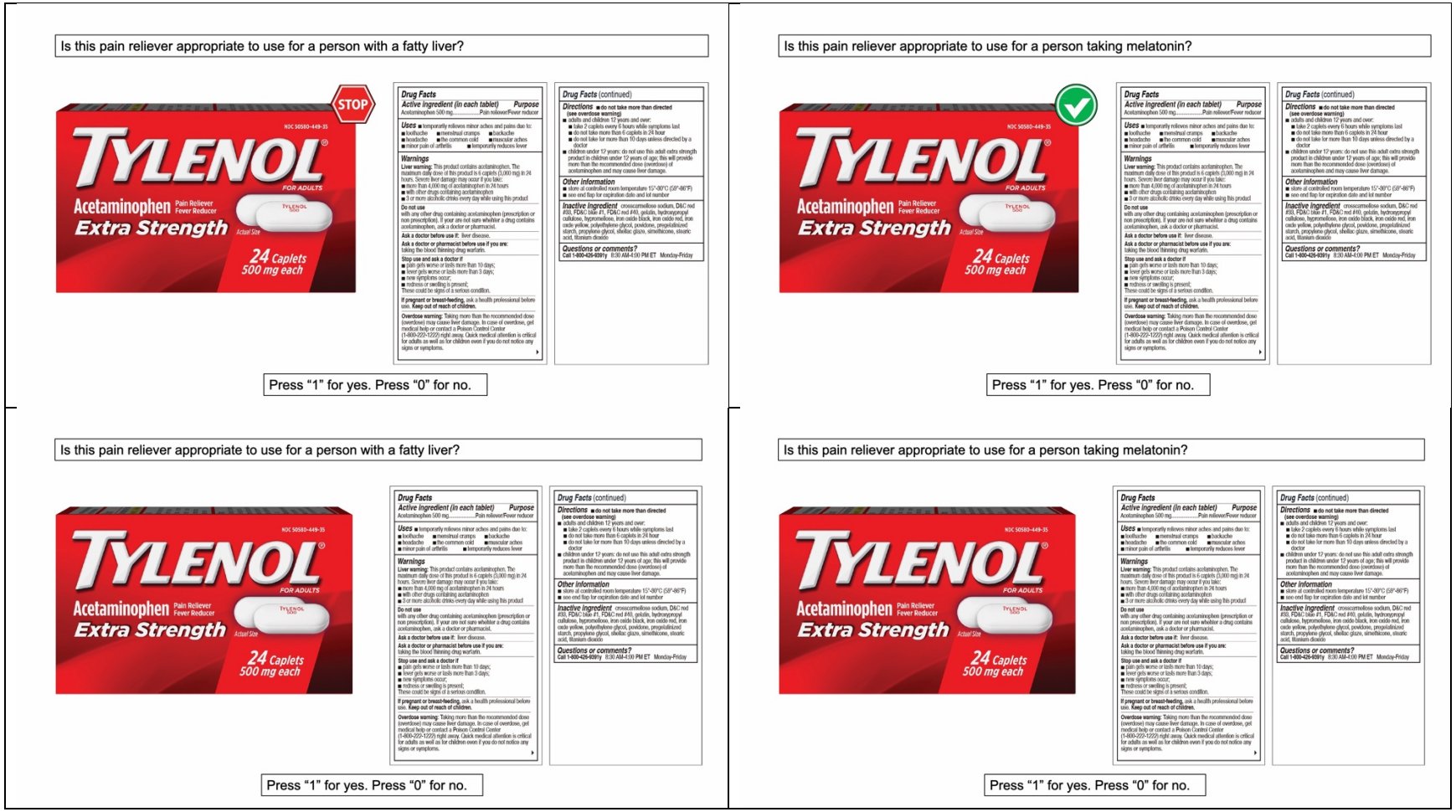


Figure X.11 The Tylenol trials of the Absolute Judgement Test

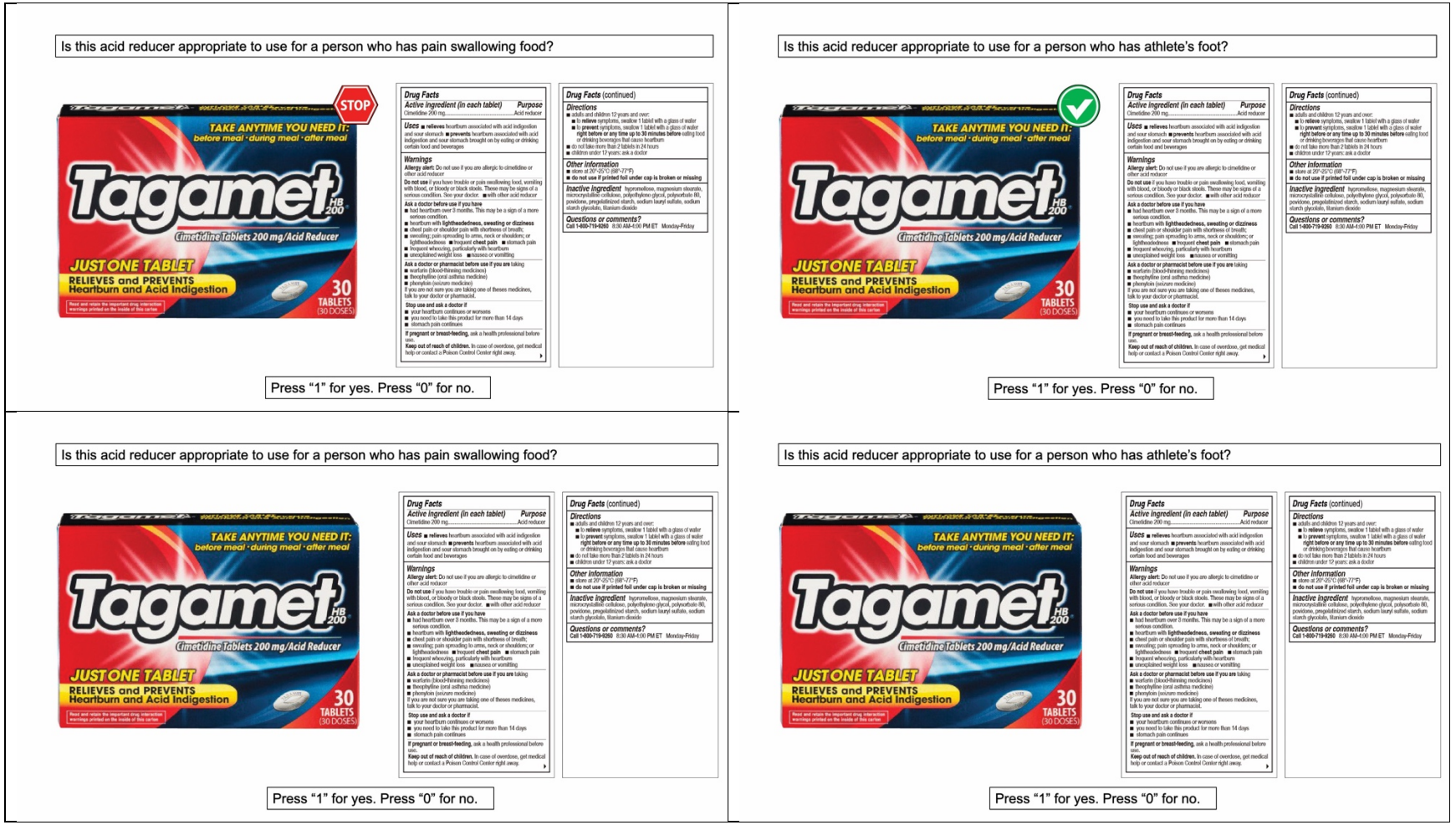


Figure X.12 The Tagamet Trials of the Absolute Judgement Test

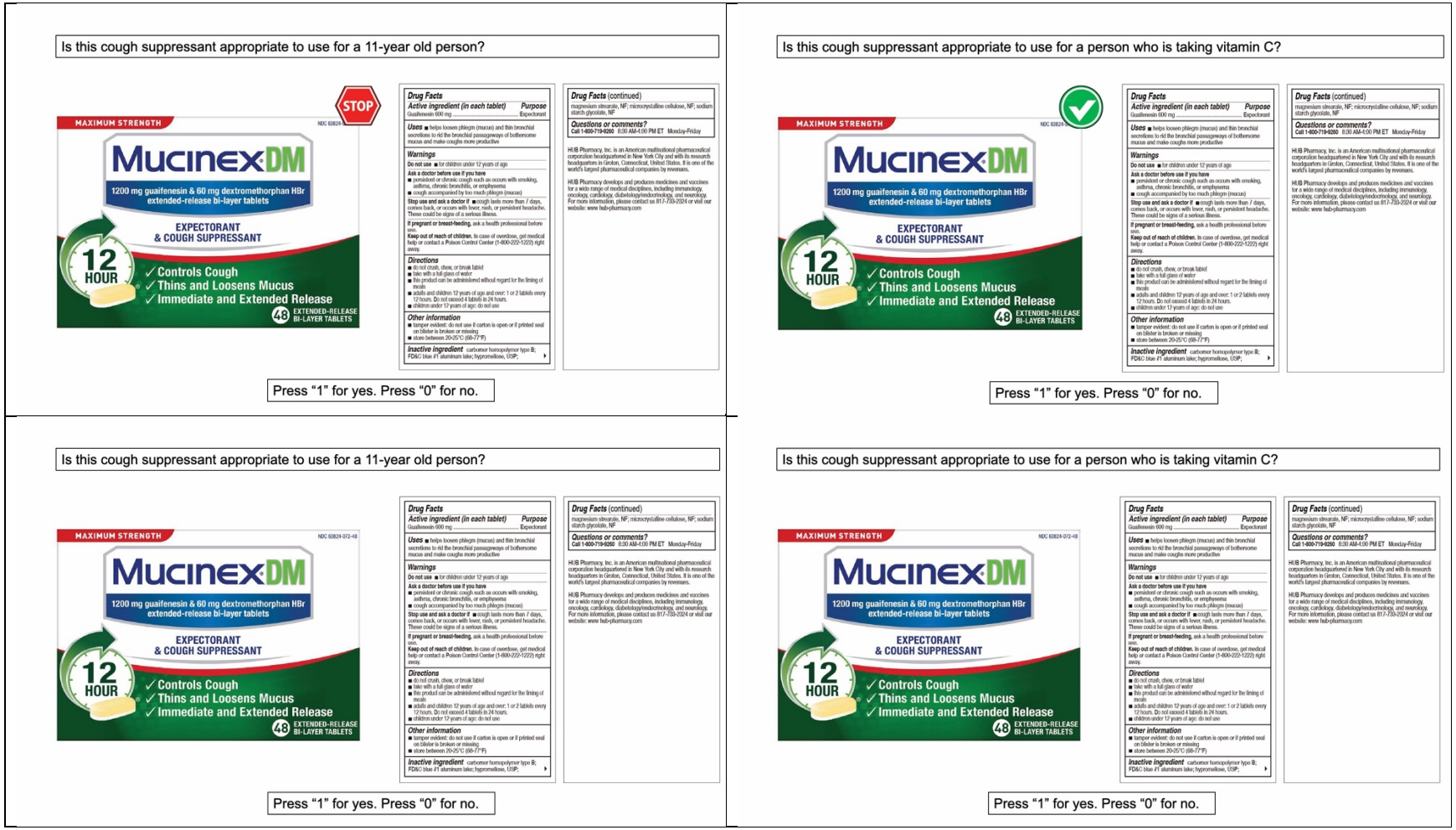


Figure X.13 The Mecinex Trials of the Absolute Judgement Test

Figure X.14 The Advil Trials of the Absolute Judgement Test

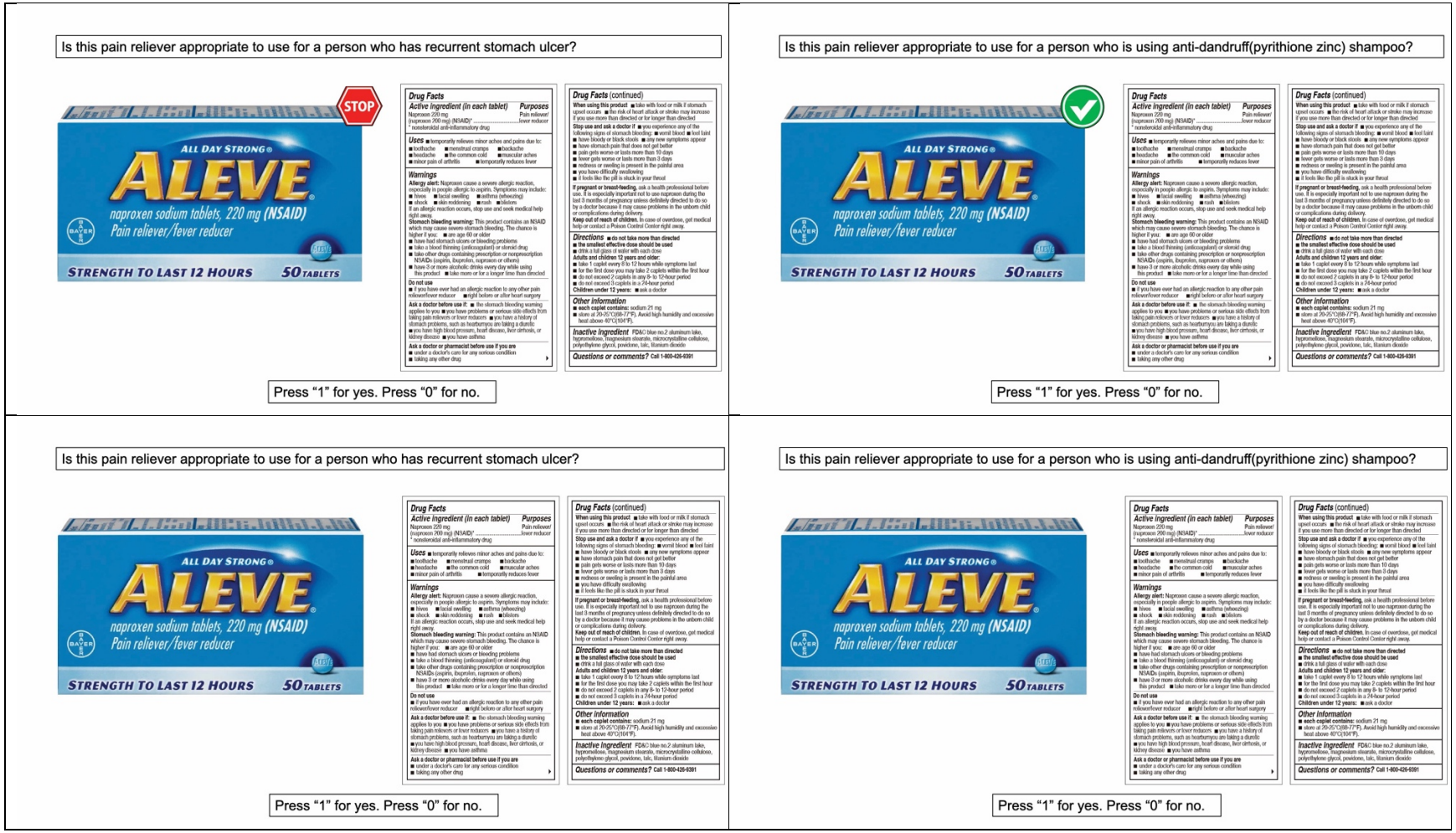


Figure X.15 The Aleve Trials of the Absolute Judgement Test

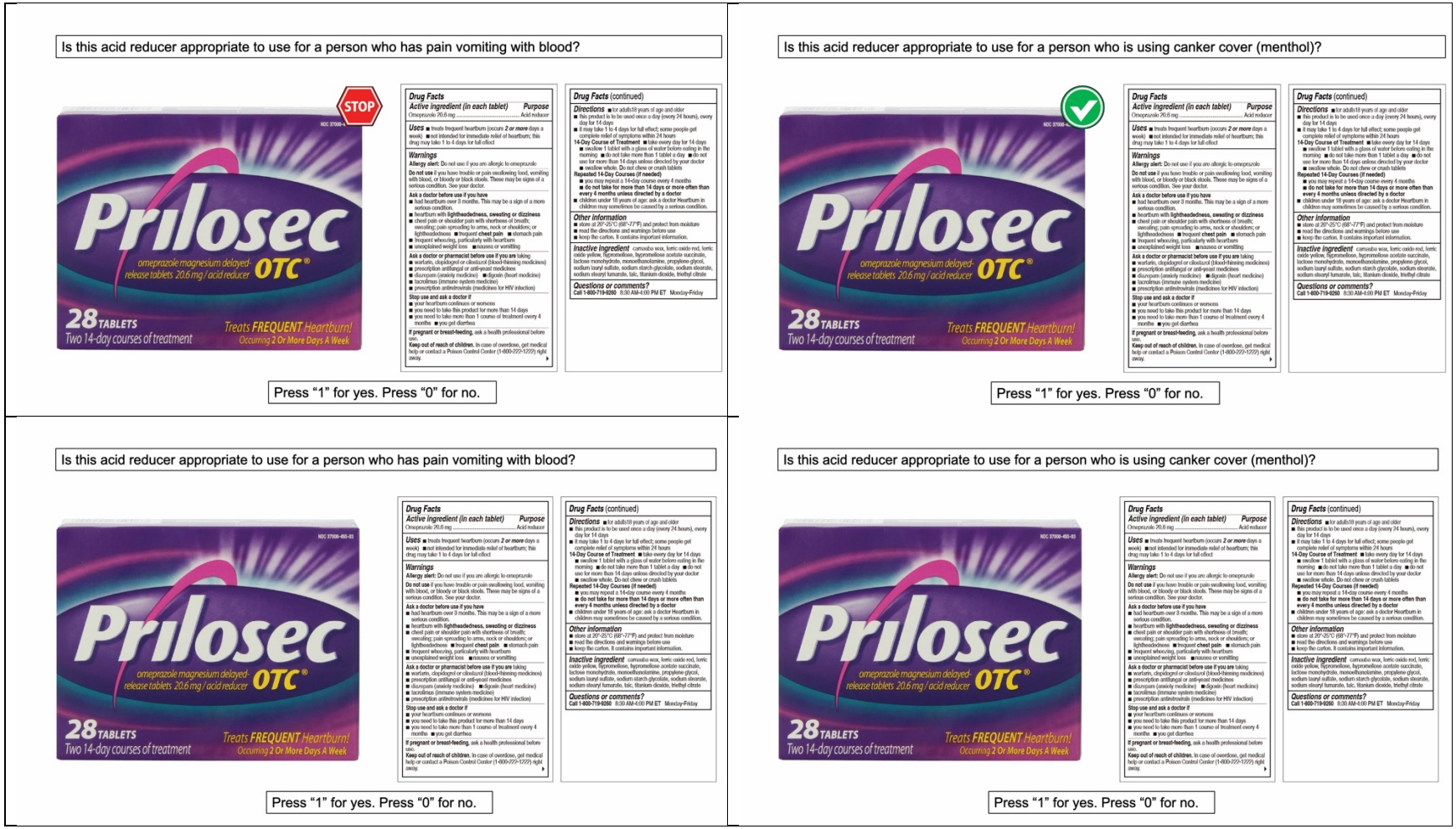


Figure X.16 The Prilosec Trials of the Absolute Judgement Test

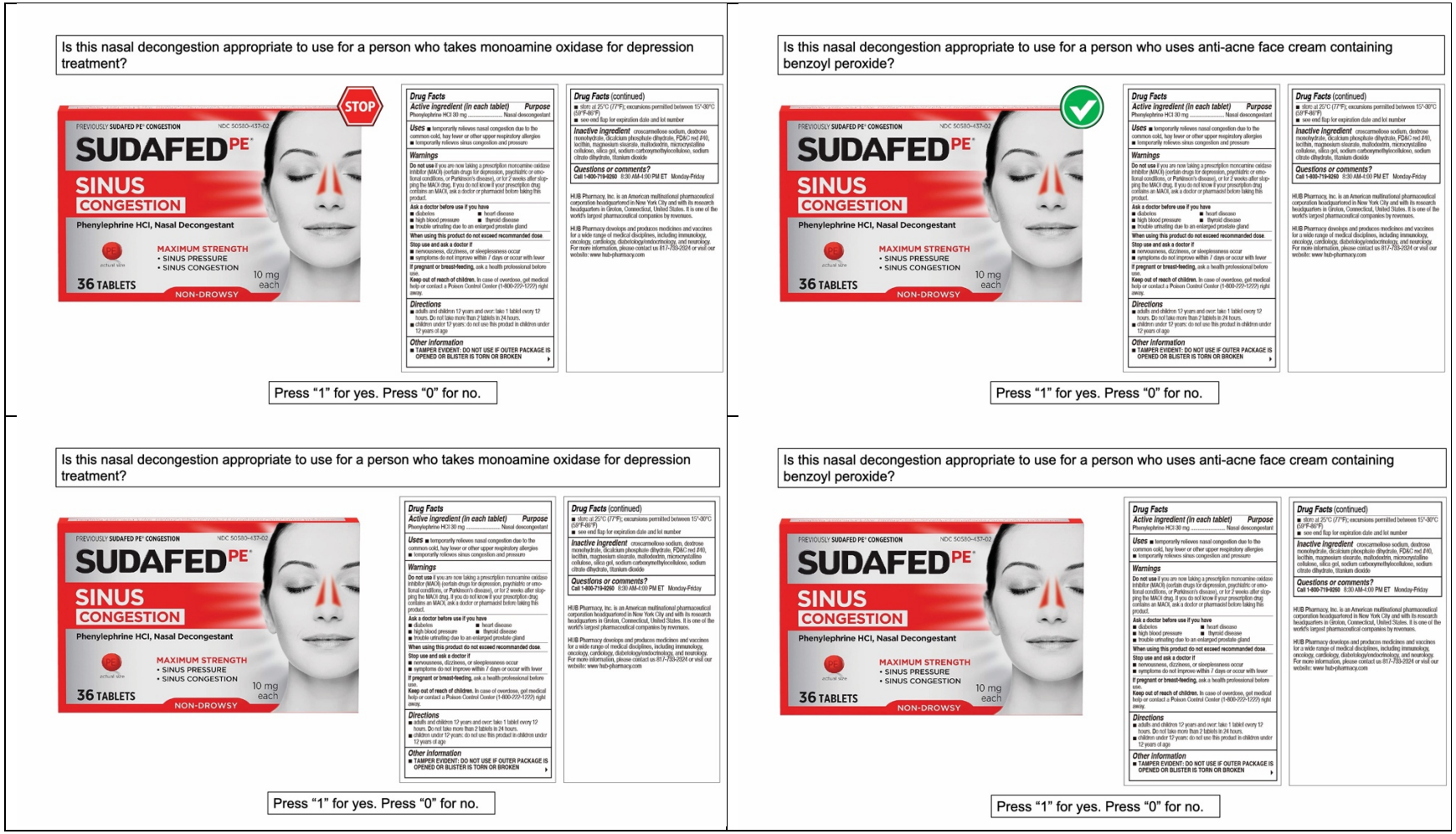


Figure X.17 The Sudafed Trials of the Absolute Judgement Test

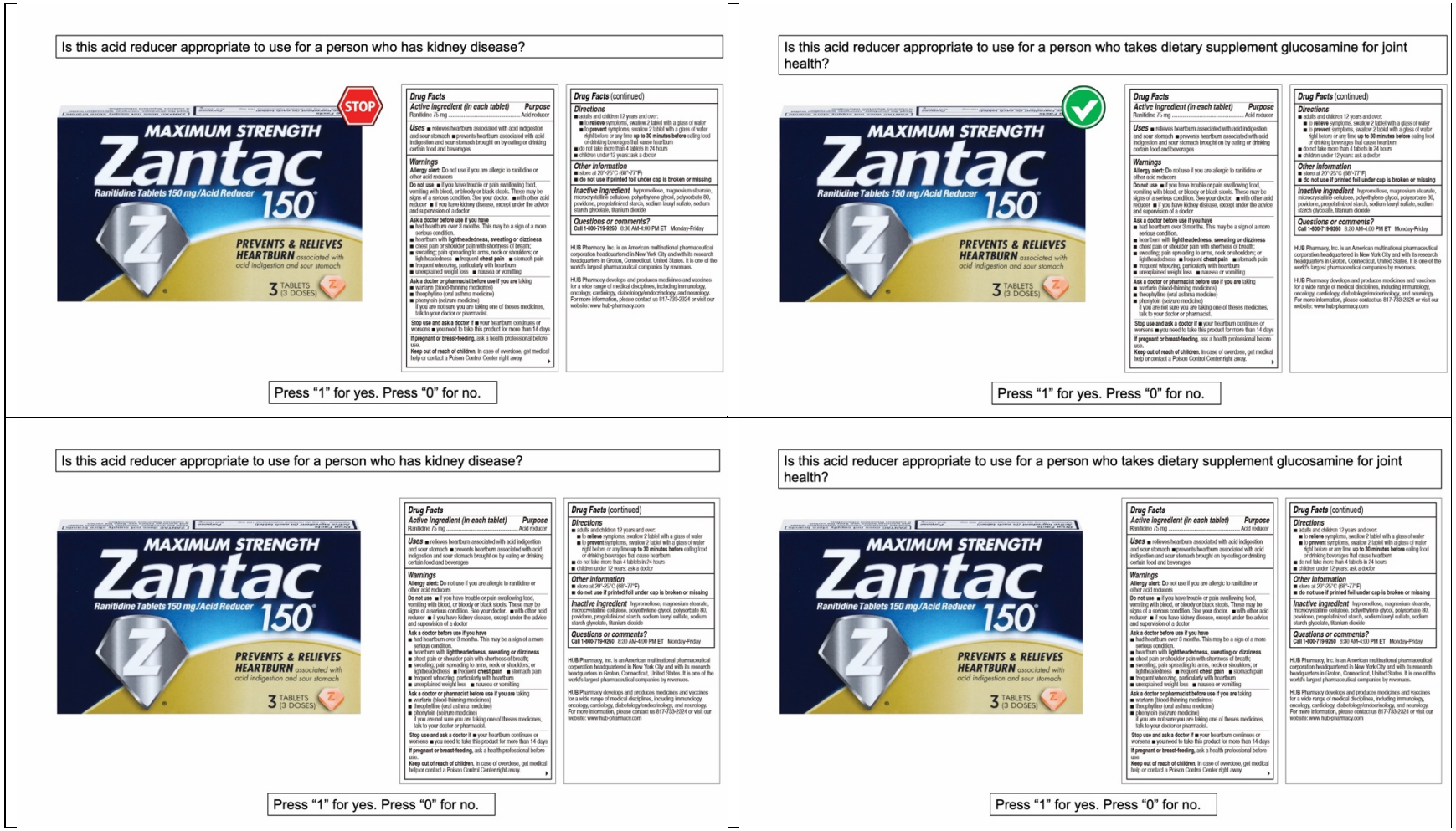
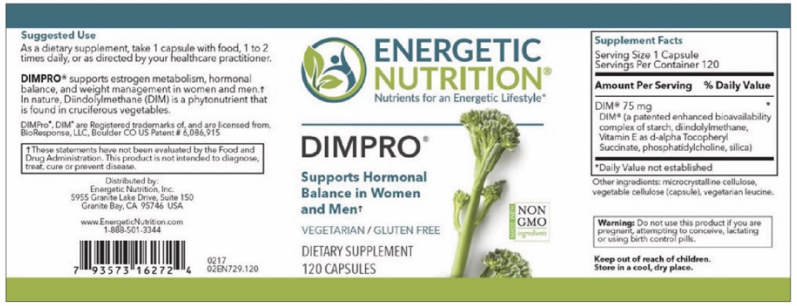


Figure X.18 The Zantac Trials of the Absolute Judgement Test

Is this dietary supplement appropriate to take for a person who has higher possibility to get prostate cancer?



Press “1” for yes. Press “0” for no.

Is this dietary supplement appropriate to take for a person who is looking for control body weight?

Press “1” for yes. Press “0” for no.

Is this dietary supplement appropriate to take for a person who wants to control the cholesterol level in body?

Press “1” for yes. Press “0” for no.

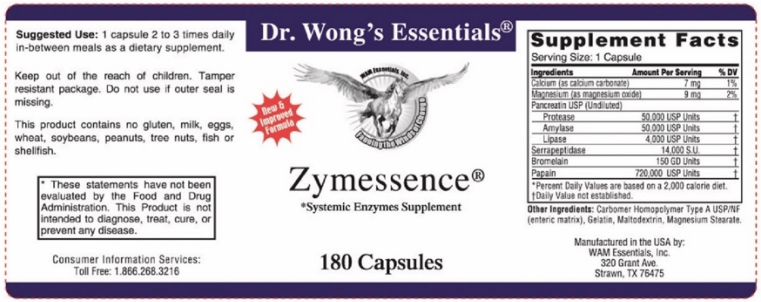
Is this dietary supplement appropriate to take for a person who has no appetite for food?

Press “1” for yes. Press “0” for no.

Figure X.19 The “Dummy” trials of the Absolute Judgement Test

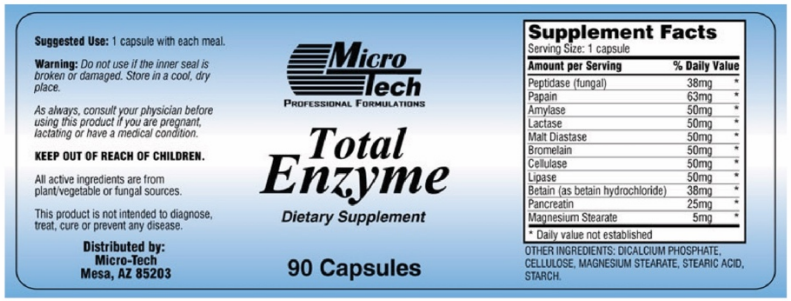
Figure X.19 (Cont'd)

Is this dietary supplement appropriate to take for a person who has no appetite for food?



Press "1" for yes. Press "0" for no.

Is this dietary supplement appropriate to take for a person who has no appetite for food?



Press "1" for yes. Press "0" for no.

APPENDIX H Flyer for Absolute Judgement Tests

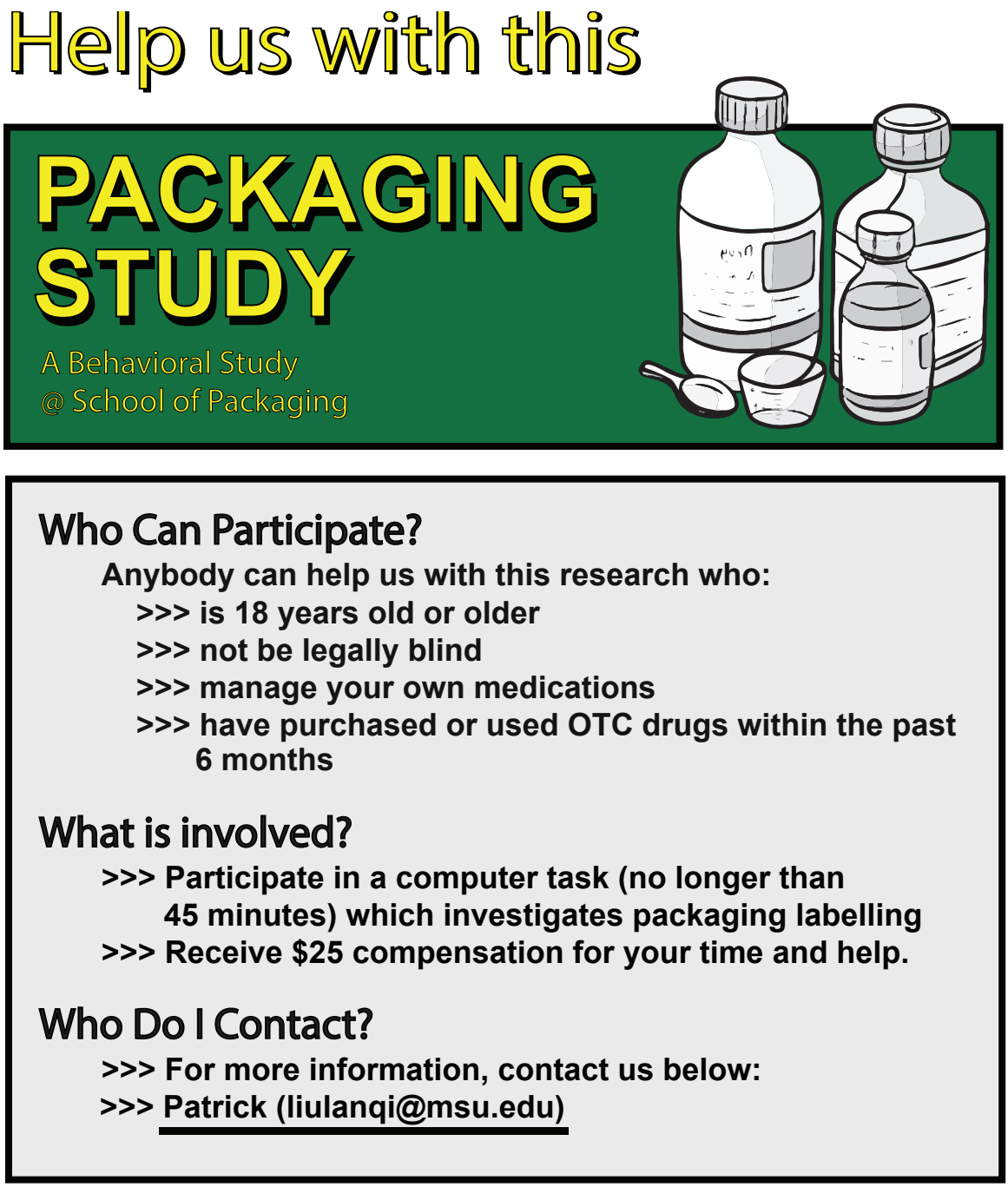


Figure X.20 Flyer for Absolute Judgement Test

APPENDIX I Consent Form for Absolute Judgement Tests

Michigan State University
School of Packaging
Study Title: Absolute Judgement Task for OTC Decision Making

INSTRUCTIONS AND RESEARCH CONSENT FORM

You are being asked to participate in a research project. Researchers are required to provide a consent form to inform you about the study, to convey that participation is voluntary, to explain risks and benefits of participation, and to empower you to make an informed decision. You should feel free to ask the researchers any questions or concerns you may have during the experiment.

1. **PURPOSE OF RESEARCH:** You are being asked to participate in an experiment to make decision of the appropriateness of OTC products under the given scenarios.

2. **TO PARTICIPATE IN THIS STUDY, YOU MUST:**
    a. Be 18 years old or older
    b. Not legally blind
    c. Have used OTC drugs during the past 6 months
    d. Have the ability to come the School of Packaging at the Michigan State University where the research was conducted.

3. WHAT YOU WILL DO: If you agree to participate in this study, the following events will take place. We will ask you to answer some basic questions about yourself. Your visual acuity, and your ability to see color will be tested; we will also test you for familiarity with medical terms. You will sit in front of a computer screen. On the computer screen, the test trials will be shown. In each trial, an OTC product image will be displayed with a brief description of a given scenario. You will then be asked to make a decision regarding the appropriateness to use the product under the given scenarios by pressing the button "1" for "Yes, it is appropriate" or "0" for "No, it is not". If you are unfamiliar, unable, or uncomfortable with using the keypad to respond, you can speak out your choice and the research team will do this for you. This process will repeat for a series of trials. The research should take no more than 45 minutes of your time. In exchange for your participation in this study, you will receive $25.

3. POTENTIAL BENEFITS: There will be no direct benefit to you from these procedures. However, it is our goal to understand what factors make certain parts of a label more noticeable than others so that we can develop labels that provide important information to people in ways that they are likely to see it.

4. POTENTIAL RISKS: We will ask you to read aloud a series of words used by medical people. It is possible that you may not be familiar with these words and this would be embarrassing. You can skip any words you are unsure of. If you are injured as a result of your participation in this research project, researchers from Michigan State University will assist you in obtaining emergency care, if necessary, for your research-related injuries. If you have insurance for medical care, your insurance carrier will be billed in the ordinary manner. As with any medical insurance, any costs that are not covered or in excess of what are paid by your insurance, including deductibles, will be your responsibility.

The University's policy is not to provide financial compensation for lost wages, disability,

pain or discomfort unless required by law to do so. This does not mean that you are giving up any legal rights you may have.

In the event that you are uncomfortable with any of the tasks, you may elect to skip a portion of the study, or discontinue altogether.

5. PRIVACY AND CONFIDENTIALITY: The data for this project will be tied to subject number, not name. Although the researchers, research staff, and the Institutional Review Board will have access to the data, neither the researchers nor anyone else will be able to link your data to you. Participant confidentiality will be protected to the maximum extent allowable by law. Paper records will be kept in Dr. Laura Bix's office for as long as required by publishers or at least three years after the study closes whichever is longer; digital records will be housed on computers in our laboratories (Packaging). Data would be provided (deidentified) to publications that deemed it a necessary part of due diligence and is also accessible to the IRB.

6. YOUR RIGHTS TO PARTICIPATE, SAY NO, OR WITHDRAW: Participation in this research is completely voluntary. Refusal to participate will involve no penalty or loss of benefit. You may also refuse to answer particular questions. You may change your mind at any time, for any reason, and withdraw without penalty or loss of compensation.

7. COSTS AND COMPENSATION FOR BEING IN THE STUDY: You will receive $25 in exchange for your participation in this study.

8. CONTACT INFORMATION FOR QUESTIONS AND CONCERNS: If you have any concerns or questions about this research study, such as scientific issues, how to do any part of it, or to report an injury, please contact the researcher (Laura Bix, Ph.D., Associate Professor of Psychology at Michigan State University: (517) xxx-xxxx, e-mail: 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.ed**Error! Hyperlink reference not valid.**

9. DOCUMENTATION OF INFORMED CONSENT: I voluntarily agree to participate in the study.

_____ _____

Signature
Date

APPENDIX J Post-test Debriefing Questions

Debriefing Questions

(After the main test, please guide the participant outside for this session.)

1. Which parts of the product package helped you answer the questions?

2. Did you notice about some design symbols on the front face of some packages? (For the concept-educated group, did you notice about personalized labeling symbols on the front face of some packages?)

3. Are the trial questions designed straightforward for you to take actions?

4. What were the most difficult trials for you to make decisions?

APPENDIX K Change Detection Accuracy for Analyses 1

K.1 Effects of Top Significant Fixed Effects

Table X.1-X.2 present statistical results of estimated marginal means and simple contrasts for the significant fixed effects of variables: highlight, change information, and ingredients. The estimated marginal means were back transformed to the original scale, which is in format of probability. For the simple contrasts within the levels of those four significant fixed effects, in comparison with basis level as mentioned in Table X.2, if the value of contrast estimate is positive, it means participants under such level for contrast were more likely (in average) to successfully detect changes compared to the basis; and if negative, then they were less likely to detect the comparative change than the base.

For the fixed effect of highlight, participants were more likely to detect changes that were highlighted (ME=0.932, SE=0.008) than those were not highlighted (ME=0.770, SE=0.019) with a difference of 16.2%. (SE=0.015, $p<0.001$).

For the fixed effects of change information, changes to the active ingredient were more likely to be detected (ME=0.935, SE=0.008) than changes to drug-diagnosis information were changing (ME=0.803, SE=0.019) ($p=5.33E-15<0.001$, contrast estimates=-0.131, SE=0.016). This was also true for changes in those with drug-drug interaction information were changing. (ME=0.842, SE=0.017) ($p=3.36E-11<0.001$, contrast estimates=-0.093, SE=0.014)

For the fixed effects of ingredients (which means the average mean of accuracy of successfully detecting changes across all critical changes for one ingredient or brand), dextromethorphan (ME=0.915, SE=0.010), ranitidine (ME=0.813, SE=0.018) or ibuprofen (ME=0.869, SE=0.014) as drug ingredients with contrast estimates -0.102 ($p=2.45E-13<0.001$, SE=0.013) and -0.046 ($p=2.01E-05<0.001$, SE=0.011) respectively.

Table X.1 Estimated Marginal Means of Top Significant Fixed Effects Change Detection Accuracy

	Mean	*Std. Error*	*95% Confidence Interval*	
			Lower	Upper
Highlight				
Highlight	0.932	0.008	0.914	0.947
Not Highlight	0.77	0.019	0.731	0.805
Change Information				
DD2 (Drug-Diagnosis)	0.803	0.019	0.763	0.839
DD1 (Drug-Drug)	0.842	0.017	0.807	0.872
AI (Active Ingredient)	0.935	0.008	0.918	0.948
Ingredient				
Ranitidine	0.813	0.018	0.776	0.846
Ibuprofen	0.869	0.014	0.84	0.894
Dextromethorphan	0.915	0.01	0.893	0.933

Continuous predictors are fixed at the following values: Age=31.55

Table X.2 Simple Contrasts of Top Significant Fixed Effects on Change Detection Accuracy

Simple Contrasts	*Contrast Estimate*	*Std. Error*	*t*	*df*	*Adj. Sig.*	*95% Confidence Interval*	
						Lower	Upper
Highlight	**(Not Highlight as basis)**						
Highlight	0.162	0.015	10.53	282	0.000	0.132	0.192
Change Information	**(Active Ingredient as basis)**						
DD2 (Drug-Diagnosis)	-0.131	0.016	-8.101	660	5.33E-15	-0.168	-0.095
DD1 (Drug-Drug)	-0.093	0.014	-6.695	1151	3.36E-11	-0.12	-0.065
Ingredient	**(Dextromethorphan as basis)**						
Ranitidine	-0.102	0.013	-7.526	935	2.45E-13	-0.132	-0.071
Ibuprofen	-0.046	0.011	-4.267	5503	2.01E-05	-0.067	-0.025

The sequential Bonferroni adjusted significance level is .05. Confidence interval bounds are approximate.

K.2 Significant 2-way interaction: Label type x Highlight

Pairwise comparisons were conducted to interpret the significant 2-way interaction that was identified between the Label types (FOP vs standard) and highlight (highlight/not highlight) (Figure X.21 and Table X.3-X.5). When label contents were highlighted, no significant differences were found in participant's change detection performance between the trials with

front-of-pack labels (ME=0.929, SE=0.008) and the trials with standard labels (ME=0.935, SE=0.011). (p=0.63>0.05) In contrast, when label contents were not highlighted, evidence were found that participants are more likely to detect changes correctly if the trials with front-of-pack labels (ME=0.804, SE=0.018) were present than those trials with standard labels (ME=0.733, SE=0.026). (p<0.001, contrast estimate=0.071, SE=0.022)

Table X.3 Results of Estimated Means of the Interaction between Label type and Highlight on Change Detection Accuracy

Label type	*Highlight*	*Mean*	*Std. Error*	*95% Wald Confidence Interval*	
				Lower	Upper
Front-of-Pack	Highlight	0.929	0.008	0.911	0.944
	Not Highlight	0.804	0.018	0.766	0.836
Standard	Highlight	0.935	0.011	0.909	0.954
	Not Highlight	0.733	0.026	0.679	0.78

Continuous predictors are fixed at the following values: Age=30.61

Table X.4 Pairwise Comparisons of Effect of Label type on Change Detection Accuracy under each Highlight level

Highlight	**Label type Pairwise Contrasts**		**Contrast Estimate**	**Std. Error**	**t**	**df**	**Adj. Sig.**	**95% Confidence Interval**	
								Lower	Upper
Highlight	Front of Pack	Standard	-0.006	0.012	-0.482	5503	6.30 E-01	-0.029	0.017
Not Highlight	Front of Pack	Standard	0.071	0.022	3.219	5503	0.001	0.028	0.114

The sequential Bonferroni adjusted significance level is .05; Confidence interval bounds are approximate.

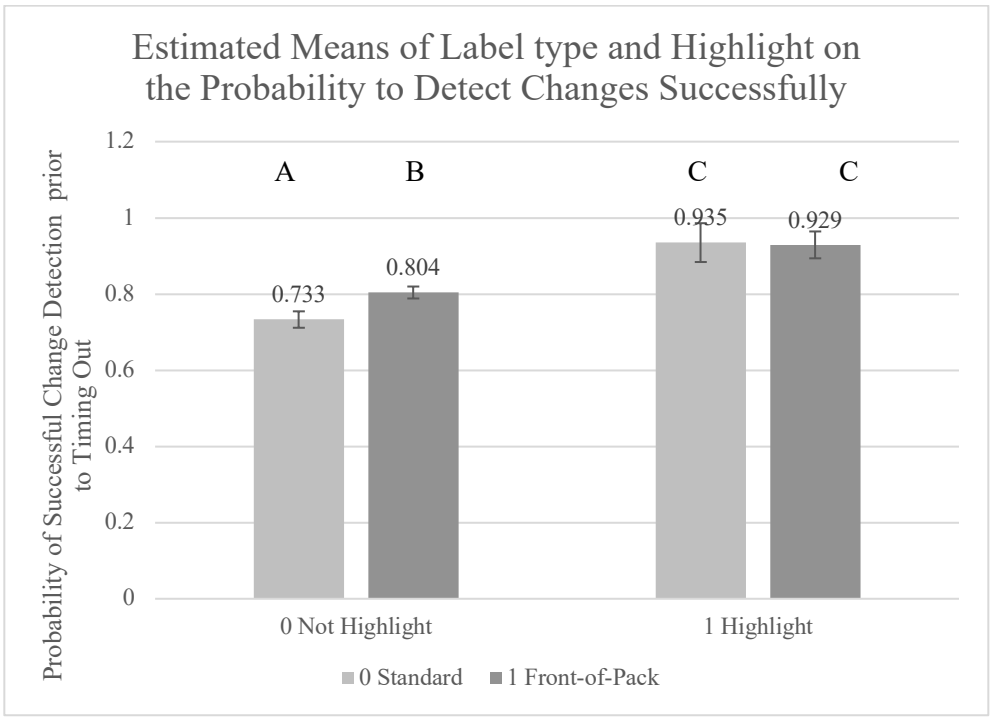


Figure X.21 Estimated Means of Label type and Highlight on the Probability to Detect Changes Successfully

Table X.5 Pairwise Comparisons of Effect of Highlight on Change Detection Accuracy under each Label type level

Label type	Highlight Pairwise Contrasts		Contrast Estimate	Std. Error	t	df	Adj. Sig.	95% Confidence Interval	
								Lower	Upper
Front-of-Pack	Highlight	Not Highlight	0.126	0.014	8.772	494	0	0.098	0.154
Standard	Highlight	Not Highlight	0.202	0.024	8.44	836	2.22E-16	0.155	0.249

The sequential Bonferroni adjusted significance level is .05; Confidence interval bounds are approximate.

APPENDIX L Change Detection Response Time for Analyses 1

L.1 Effects of Top Significant Fixed Effects

Table X.6-X.7 present statistical results of estimated marginal means and simple contrasts for the significant fixed effects of variables as aforementioned: highlight, change information, and ingredients. The estimated marginal means were back transformed to the original scale from the format of natural log. For the simple contrasts within the levels of those five significant fixed effects, in comparison with basis level as mentioned in Table X.7, if the value of contrast estimate is indicated as positive, it means participants on average, took longer time to correctly detect changes than for those trials that served as the base; and if negative, then it took shorter time in average.

Table X.6 Estimated Marginal Means of Top Significant Fixed Effects on Response Time for Correct Change Detections

	Mean	*Std. Error*	*95% Confidence Interval*	
			Lower	Upper
Label type				
Front-of-Pack	4.542	0.238	4.099	5.034
Standard	4.233	0.225	3.814	4.699
Highlight				
Highlight	3.79	0.2	3.418	4.203
Not Highlight	5.073	0.268	4.574	5.628
Change Location				
Principal Display Panel	3.029	0.161	2.729	3.362
Drug Facts Label	6.348	0.333	5.728	7.035
Change Information				
DD2 - Drug-Diagnosis	5.244	0.28	4.722	5.823
DD1 - Drug-Drug	4.769	0.255	4.295	5.295
AI - Active Ingredients	3.372	0.178	3.041	3.739
Ingredients				
Ranitidine	4.652	0.247	4.192	5.163
Ibuprofen	4.294	0.228	3.87	4.765
Dextromethorphan	4.221	0.224	3.805	4.683

Continuous predictors are fixed at the following values: Age=30.61

For all the trials from the treatments which included FOPs, participants took longer, on average, to correctly detect changes (ME=4.542, SE=0.238) than the trials with standard labels

(ME=4.233, SE=0.225) (contrast estimate=0.309, SE=0.065, p=1.76E-06<0.001). Participants took significantly less time to correctly detect changes with the content highlighted (ME=3.790, SE=0.200) than those were not highlighted (ME=5.073, SE=0.268) (contrast estimate=-1.283, SE=0.091, p<0.001) on average across the other treatment effects. For the fixed effect of change location, participants responded significantly faster for the trials with changes happened on principal display panel (PDP) (ME=3.029, SE=0.161) than those on drug facts labels (DFL) (ME=6.348, SE=0.333). (p<0.001, contrast estimates=-3.318, SE=0.183) When results for change information was collapsed across conditions, it took participants significantly less time to detect changes to the active ingredients (ME=3.372, SE=0.178) than changes to drug-diagnosis information (ME=5.244, SE=0.28) (p<0.001, contrast estimates=1.872, SE=0.123); this was also true of drug-drug interaction information were changing. (ME=4.769, SE=0.255) (p<0.001, contrast estimates1.397, SE=0.064)

Table X.7 Simple Contrasts of Top Significant Fixed Effects on Response Time for Correct Change Detections

	Contrast Estimate	*Std. Error*	*t*	*df*	*Adj. Sig.*	*95% Confidence Interval*	
						Lower	Upper
FOPtype	**Standard as comparing base**						
Front-of-Pack	0.309	0.065	4.785	4632	1.76E-06	0.183	0.436
Highlight	**Not Highlight as comparing base**						
Highlight	-1.283	0.091	-14.034	4632	0	-1.462	-1.104
Change Location	**Drug Facts Label as comparing basis**						
Principal Display Panel	-3.318	0.183	-18.148	4632	0	-3.677	-2.96
Change Information	**AI (Active Ingredient) as comparing base**						
DD2 - Drug-Diagnosis	1.872	0.123	15.275	4632	0	1.597	2.147
DD1 - Drug-Drug	1.397	0.1	14.029	4632	0	1.202	1.592
Ingredients	**Dextromethorphan as comparing base**						
Ranitidine	0.431	0.071	6.044	4632	3.24E-09	0.271	0.591
Ibuprofen	0.073	0.064	1.141	4632	0.254	-0.052	0.197

The sequential Bonferroni adjusted significance level is .05; Confidence interval bounds are approximate.

Using the reaction time to detect changes to the dextromethorphan brand as a basis for comparison (ME=4.221, SE=0.224), participants took significantly longer time to correctly detect changes in the trials with ranitidine (ME=4.652, SE=0.247; p=3.24E-09<0.001, contrast estimate=0.431, SE=0.071); but there was no evidence that participants performed significant differently in the trials with ibuprofen (ME=4.294, SE=0.228; p=0.254).

L.2 Significant 2-way interaction: Label type x Highlight

Pairwise comparisons were conducted to interpret the significant 2-way interaction that was identified between the Label types (FOP vs standard) and highlight (highlight/not highlight) (Figure X.22 and Table X.8-X.10). Evidence suggested that highlighting the content shortened the time to spent to correctly detect changes when compared with trials which were not highlighted; this was true for both the trials with front-of-pack labels (p=<0.001, contrast estimate=-1.157, SE=0.096), or with standard labels. (p=<0.001, contrast estimate=-1.399, SE=0.123)

Table X.8 Results of Estimated Means of the Interaction between Label type and Highlight on Response Time for Correct Change Detection

FOP type	*Highlight*	*Mean*	*Std. Error*	*95% Confidence Interval*	
				Lower	Upper
Front-of-Pack	Highlight	4.001	0.212	3.606	4.438
	Not Highlight	5.158	0.274	4.647	5.724
Standard	Highlight	3.591	0.194	3.229	3.993
	Not Highlight	4.99	0.273	4.483	5.554

Continuous predictors are fixed at the following values: Age=30.61

When we examined the effect of change location, when label contents were highlighted, significant differences were found suggesting that participants took more time to correctly detect changes for the trials with front-of-pack labels (ME=4.001, SE=0.212) than the trials with standard labels (ME=3.591, SE=0.194). (p=3.71E-08 <0.001, contrast estimate=0.41, SE=0.074)

However, when label contents were not highlighted, no evidence was found regarding the response time to correctly detect changes between the trials with front-of-pack labels (ME=5.158, SE=0.274) and the trials with standard labels (ME=0.4.99, SE=0.273). (p=0.112)

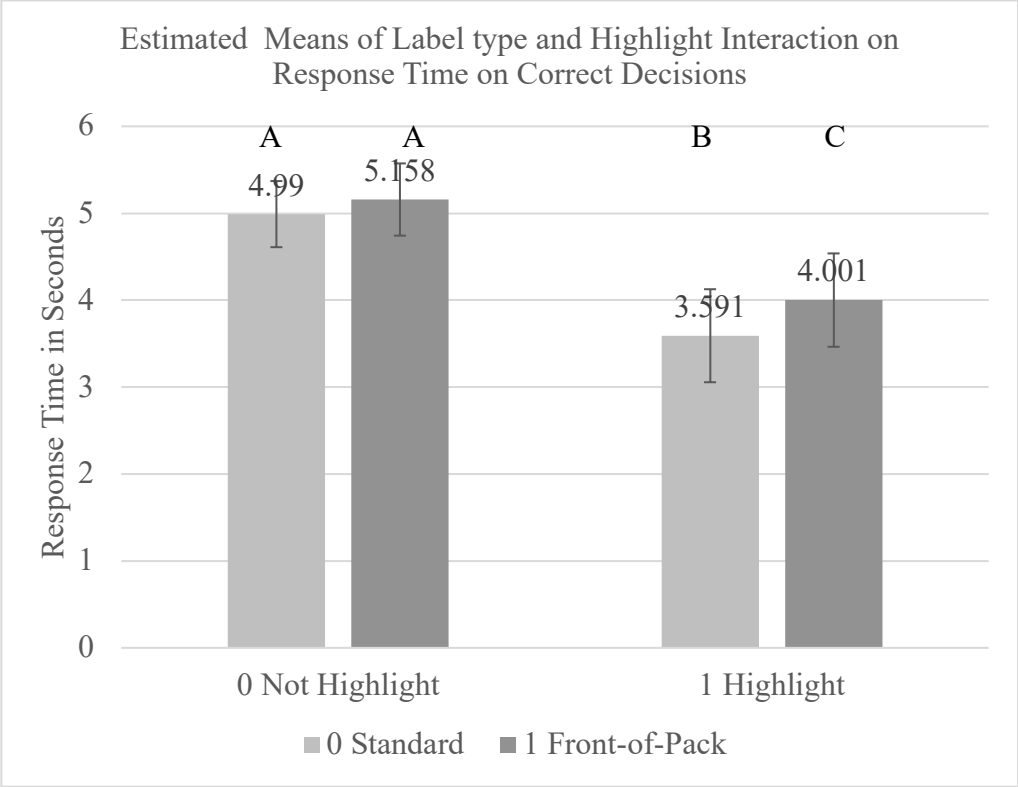


Figure X.22 Estimated Means of Label type and Highlight Interaction on Response Time on Correct Decisions

Table X.9 Pairwise Comparisons of Effect of Label type on Response Time under each Highlight level.

Highlight	**Label type Pairwise Contrasts**		**Contrast Estimate**	**Std. Error**	**t**	**df**	**Adj. Sig.**	**95% Confidence Interval**	
								Lower	Upper
Highlight	Front of Pack	Standard	0.41	0.074	5.514	4632	3.71E-08	0.264	0.555
Not Highlight	Front of Pack	Standard	0.168	0.105	1.592	4632	0.112	-0.039	0.374
The sequential Bonferroni adjusted significance level is .05. Confidence interval bounds are approximate.									

Table X.10 Pairwise Comparisons of Effect of Highlight on Response Time under each Label type level

Label type	Highlight Pairwise Contrasts		Contrast Estimate	Std. Error	t	df	Adj. Sig.	95% Confidence Interval	
								Lower	Upper
Front-of-Pack	Highlight	Not Highlight	-1.157	0.096	-12.066	4632	0	-1.345	-0.969
Standard	Highlight	Not Highlight	-1.399	0.123	-11.39	4632	0	-1.64	-1.158

The sequential Bonferroni adjusted significance level is .05; Confidence interval bounds are approximate.

BIBLIOGRAPHY

BIBLIOGRAPHY

Aker, J., Beck, M., Travis, S., & Harris, J. (2014). Consumer navigation and selection behaviors for OTC products in a retail setting. *Indianapolis, Indiana: The Consumer Healthcare Products Association*.

Aronson, J. K., & Ferner, R. E. (2005). Clarification of Terminology in Drug Safety. *Drug Safety*, *28*(10), 851–870. https://doi.org/10.2165/00002018-200528100-00003

Bennadi, D. (2013). Self-medication: A current challenge. *Journal of Basic and Clinical Pharmacy*, *5*(1), 19–23. https://doi.org/10.4103/0976-0105.128253

Berman, A. (2004). Reducing Medication Errors Through Naming, Labeling, and Packaging. *Journal of Medical Systems*, *28*(1), 9–29. https://doi.org/10.1023/B:JOMS.0000021518.60670.10

Berner, E. S. (2007). *Clinical decision support systems* (Vol. 233). Springer.

Bhansali, A. H., Sangani, D. S., Mhatre, S. K., & Sansgiry, S. S. (2018). Effect of warning placement on the information processing of college students reading an OTC drug facts panel. *Journal of American College Health*, *66*(4), 237–245. https://doi.org/10.1080/07448481.2018.1431891

Bin, S., Masood, S., & Jung, Y. (2020). Virtual and augmented reality in medicine. In *Biomedical Information Technology* (pp. 673–686). Elsevier.

Bix, L., Bello, N. M., Auras, R., Ranger, J., & Lapinski, M. K. (2009). Examining the conspicuousness and prominence of two required warnings on OTC pain relievers. *Proceedings of the National Academy of Sciences*, *106*(16), 6550–6555. https://doi.org/10.1073/pnas.0810665106

Bix, L., Seo, D. C., Ladoni, M., Brunk, E., & Becker, M. W. (2016). Evaluating Varied Label Designs for Use with Medical Devices: Optimized Labels Outperform Existing Labels in the Correct Selection of Devices and Time to Select. *PLOS ONE*, *11*(11), e0165002. https://doi.org/10.1371/journal.pone.0165002

Bix, Laura, Sundar, R. P., Bello, N. M., Peltier, C., Weatherspoon, L. J., & Becker, M. W. (2015). To See or Not to See: Do Front of Pack Nutrition Labels Affect Attention to Overall Nutrition Information? *PLOS ONE*, 20.

Cabrera, M., Machín, L., Arrúa, A., Antúnez, L., Curutchet, M. R., Giménez, A., & Ares, G. (2017). Nutrition warnings as front-of-pack labels: Influence of design features on healthfulness perception and attentional capture. *Public Health Nutrition*, *20*(18), 3360–3371.

Card, S. K. (2018). *The Psychology of Human-Computer Interaction*. CRC Press. https://doi.org/10.1201/9780203736166

Card, S., MORAN, T., & Newell, A. (1986). The model human processor- An engineering model of human performance. *Handbook of Perception and Human Performance.*, *2*(45–1).

CFR - Code of Federal Regulations Title 21. Retrieved March 18, 2021, from https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=330&showFR=1

Coleman, J. J., & Pontefract, S. K. (2016). Adverse drug reactions. *Clinical Medicine*, *16*(5), 481–485. https://doi.org/10.7861/clinmedicine.16-5-481

Cowburn, G., & Stockley, L. (2005). Consumer understanding and use of nutrition labelling: A systematic review. *Public Health Nutrition*, *8*(1), 21–28.

Cryer, B., Barnett, M. A., Wagner, J., & Wilcox, C. M. (2016). Overuse and misperceptions of nonsteroidal anti-inflammatory drugs in the United States. *The American Journal of the Medical Sciences*, *352*(5), 472–480.

de la Fuente, J., Gustafson, S., Twomey, C., & Bix, L. (2015). An affordance-based methodology for package design. *Packaging Technology and Science*, *28*(2), 157–171.

Diodati, G., Gómez, A., Martínez, M., & Luna, D. (2015). Augmented Reality: Real-Time Information Concerning Medication Consumed by a Patient. *Studies in Health Technology and Informatics*, *216*, 896–896.

Esfahanian, S., & Link to external site, this link will open in a new window. (2020). *A Patient-Centered Approach to Labeling for Over-The-Counter Medications: Using Data to Drive Design Decisions for the Benefit of Older Adults* [M.Sc., Michigan State University]. https://search.proquest.com/docview/2450152074/abstract/E12BC5F91DA14703PQ/1

FDA. (2020, May 7). *Drug Applications for Over-the-Counter (OTC) Drugs*. FDA; FDA. https://www.fda.gov/drugs/types-applications/drug-applications-over-counter-otc-drugs

Feteira-Santos, R., Fernandes, J., Virgolino, A., Alarcão, V., Sena, C., Vieira, C. P., Gregório, M. J., Nogueira, P., Costa, A., & Graça, P. (2020). Effectiveness of interpretive front-of-pack nutritional labelling schemes on the promotion of healthier food choices: A systematic review. *International Journal of Evidence-Based Healthcare*, *18*(1), 24–37.

Feunekes, G. I. J., Gortemaker, I. A., Willems, A. A., Lion, R., & van den Kommer, M. (2008). Front-of-pack nutrition labelling: Testing effectiveness of different nutrition labelling formats front-of-pack in four European countries. *Appetite*, *50*(1), 57–70. https://doi.org/10.1016/j.appet.2007.05.009

Gawasane, A., Bix, L., de la Fuente, J., Sundar, R. P., & Smith, T. J. (2012). Consumer attention to an over-the-counter warning in four different styles of design. *Packaging Technology and Science*, *25*(7), 385–396.

Guidance for Industry Labeling OTC Human Drug Products. (2009). U.S Food Drug Administration.

Hawley, D. F., & Leasure, J. L. (2012). Region-specific response of the hippocampus to chronic unpredictable stress. *Hippocampus*, *22*(6), 1338–1349.

Helmons, P. J., Suijkerbuijk, B. O., Nannan Panday, P. V., & Kosterink, J. G. W. (2015). Drug-drug interaction checking assisted by clinical decision support: A return on investment analysis. *Journal of the American Medical Informatics Association: JAMIA*, *22*(4), 764–772. https://doi.org/10.1093/jamia/ocu010

Hersey, J. C., Wohlgenant, K. C., Arsenault, J. E., Kosa, K. M., & Muth, M. K. (2013). Effects of front-of-package and shelf nutrition labeling systems on consumers. *Nutrition Reviews*, *71*(1), 1–14.

Hodgkins, C., Barnett, J., Wasowicz-Kirylo, G., Stysko-Kunkowska, M., Gulcan, Y., Kustepeli, Y., Akgungor, S., Chryssochoidis, G., Fernández-Celemin, L., Storcksdieck genannt Bonsmann, S., Gibbs, M., & Raats, M. (2012). Understanding how consumers categorise nutritional labels: A consumer derived typology for front-of-pack nutrition labelling. *Appetite*, *59*(3), 806–817. https://doi.org/10.1016/j.appet.2012.08.014

Holden, R. J., Srinivas, P., Campbell, N. L., Clark, D. O., Bodke, K. S., Hong, Y., Boustani, M. A., Ferguson, D., & Callahan, C. M. (2019). Understanding older adults' medication decision making and behavior: A study on over-the-counter (OTC) anticholinergic medications. *Research in Social and Administrative Pharmacy*, *15*(1), 53–60. https://doi.org/10.1016/j.sapharm.2018.03.002

Höllerer, T., & Feiner, S. (2004). Mobile augmented reality. *Telegeoinformatics: Location-Based Computing and Services*, *21*.

Ikonen, I., Sotgiu, F., Aydinli, A., & Verlegh, P. W. (2020). Consumer effects of front-of-package nutrition labeling: An interdisciplinary meta-analysis. *Journal of the Academy of Marketing Science*, *48*(3), 360–383.

King, J. P., Davis, T. C., Bailey, S. C., Jacobson, K. L., Hedlund, L. A., Di Francesco, L., Parker, R. M., & Wolf, M. S. (2011). Developing consumer-centered, nonprescription drug labeling: A study in acetaminophen. *American Journal of Preventive Medicine*, *40*(6), 593–598.

Liu, L. (2016). *The effect of labeling content and prominence on information processing among older adults during self-selection of over-the-counter medications*. https://etd.lib.msu.edu/islandora/object/etd:3967

Mahapatra, T. (2017). Self-care and self-medication: A commentary. *Annals of Tropical Medicine and Public Health*, *10*(3), 505. https://doi.org/10.4103/1755-6783.213160

Melville, S. (2021, January 4). Pharmacy Outlook: Scott Melville, CHPA. *CDR – Chain Drug Review*. https://www.chaindrugreview.com/pharmacy-outlook-scott-melville-chpa-2/

Monk, A. (1999). Modelling cyclic interaction. *Behaviour & Information Technology*, *18*(2), 127–139.

Moro, C., Štromberga, Z., Raikos, A., & Stirling, A. (2017). The effectiveness of virtual and augmented reality in health sciences and medical anatomy. *Anatomical Sciences Education*, *10*(6), 549–559.

Nijman, C. A. J., Zijp, I. M., Sierksma, A., Roodenburg, A. J. C., Leenen, R., van den Kerkhoff, C., Weststrate, J. A., & Meijer, G. W. (2007). A method to improve the nutritional quality of foods and beverages based on dietary recommendations. *European Journal of Clinical Nutrition*, *61*(4), 461–471. https://doi.org/10.1038/sj.ejcn.1602548

OTC Retail Sales 1964-2019. (n.d.). Retrieved April 25, 2020, from https://www.chpa.org/OTCRetailSales.aspx

Parekh, P., Patel, S., Patel, N., & Shah, M. (2020). Systematic review and meta-analysis of augmented reality in medicine, retail, and games. *Visual Computing for Industry, Biomedicine, and Art*, *3*(1), 21. https://doi.org/10.1186/s42492-020-00057-7

Pelizaro, C., & Mcdonald, D. (2006). Spatial (GIS-based) decision support system for the Westernport region. *Applied Gis*, *2*. https://doi.org/10.2104/ag060017

Rensink, R. A. (2002). Change Detection. *Annual Review of Psychology*, *53*(1), 245–277. https://doi.org/10.1146/annurev.psych.53.100901.135125

Rojas, K. M., & Li, H. (2017). Adverse Events and Over-the-Counter (OTC) Drugs: Is Inappropriate Labeling the Problem? - The Case of Acetaminophen. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *61*(1), 676–680. https://doi.org/10.1177/1541931213601656

Roseman, M. G., Joung, H.-W., & Littlejohn, E. I. (2018). Attitude and Behavior Factors Associated with Front-of-Package Label Use with Label Users Making Accurate Product Nutrition Assessments. *Journal of the Academy of Nutrition and Dietetics*, *118*(5), 904–912. https://doi.org/10.1016/j.jand.2017.09.006

Schmiedl, S., Rottenkolber, M., Hasford, J., Rottenkolber, D., Farker, K., Drewelow, B., Hippius, M., Saljé, K., & Thürmann, P. (2014). Self-Medication with Over-the-Counter and Prescribed Drugs Causing Adverse-Drug-Reaction-Related Hospital Admissions: Results of a Prospective, Long-Term Multi-Centre Study. *Drug Safety*, *37*(4), 225–235. https://doi.org/10.1007/s40264-014-0141-3

Shackel, B., & Richardson, S. J. (1991). *Human factors for informatics usability*. Cambridge university press.

Shaver, E. F., & Wogalter, M. S. (2003). A Comparison of Older vs. Newer Over-the-Counter (OTC) Nonprescription Drug Labels on Search Time Accuracy. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *47*(5), 826–830. https://doi.org/10.1177/154193120304700508

Shehnaz, S. I., Agarwal, A. K., & Khan, N. (2014). A Systematic Review of Self-Medication Practices Among Adolescents. *Journal of Adolescent Health*, *55*(4), 467–483. https://doi.org/10.1016/j.jadohealth.2014.07.001

Shiffman, S., Cotton, H., Jessurun, C., Sembower, M. A., Pype, S., & Phillips, J. (2016). Testing of candidate icons to identify acetaminophen-containing medicines. *Pharmacy*, *4*(1), 10.

Sprague, R. H. (1980). A Framework for the Development of Decision Support Systems. *MIS Quarterly*, *4*(4), 1–26. https://doi.org/10.2307/248957

Sutherland, J., Belec, J., Sheikh, A., Chepelev, L., Althobaity, W., Chow, B. J., Mitsouras, D., Christensen, A., Rybicki, F. J., & La Russa, D. J. (2019). Applying modern virtual and augmented reality technologies to medical images and models. *Journal of Digital Imaging*, *32*(1), 38–53.

Sutton, R. T., Pincock, D., Baumgart, D. C., Sadowski, D. C., Fedorak, R. N., & Kroeker, K. I. (2020). An overview of clinical decision support systems: Benefits, risks, and strategies for success. *Npj Digital Medicine*, *3*(1), 17. https://doi.org/10.1038/s41746-020-0221-y

Takemoto, J. K., Bratelli, R. A., Parmentier, B. L., Merritt, T. A., & Coyne, L. (2019). Extended Reality in Patient Care and Pharmacy Practice: A Viewpoint. *Journal of Contemporary Pharmacy Practice*, *66*(4), 33–38. https://doi.org/10.37901/jcphp18-00030

Tong, V., Raynor, D. K., & Aslani, P. (2014). Design and comprehensibility of over-the-counter product labels and leaflets: A narrative review. *International Journal of Clinical Pharmacy*, *36*(5), 865–872. https://doi.org/10.1007/s11096-014-9975-0

Tong, V., Raynor, D. K., & Aslani, P. (2017). User testing as a method for identifying how consumers say they would act on information related to over-the-counter medicines. *Research in Social and Administrative Pharmacy*, *13*(3), 476–484. https://doi.org/10.1016/j.sapharm.2016.06.001

Tran, H. H., Suenaga, H., Kuwana, K., Masamune, K., Dohi, T., Nakajima, S., & Liao, H. (2011). Augmented reality system for oral surgery using 3D auto stereoscopic visualization. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 81–88.

Trivedi, H., Trivedi, A., & Hannan, M. F. (2014). Readability and comprehensibility of over-the-counter medication labels. *Renal Failure*, *36*(3), 473–477. https://doi.org/10.3109/0886022X.2013.872571

Westerlund, T., Barzi, S., Bernsten, C., Westerlund, T., Barzi, S., & Bernsten, C. (2017). Consumer views on safety of over-the-counter drugs, preferred retailers and information sources in Sweden: After re-regulation of the pharmacy market. *Pharmacy Practice (Granada)*, *15*(1). https://doi.org/10.18549/pharmpract.2017.01.894

Zhao, Y., & Ma, S. (2016). Observations on the Prevalence, Characteristics, and Effects of Self-Treatment. *Frontiers in Public Health*, *4*. https://doi.org/10.3389/fpubh.2016.00069