Toward a Computational Model and Decision Support System for Reducing Errors in Pharmaceutical Packaging

Carson Quigley
Biomedical Engineering
Bucknell University
17837, Lewisburg, PA
carson.quigley@bucknell.edu

Steven Shooter
Mechanical Engineering
Bucknell University
17837, Lewisburg, PA
shooter@bucknell.edu

Aaron Mitchel
Psychology
Bucknell University
17837, Lewisburg, PA
aaron.mitchel@bucknell.edu

Scarlett Miller
School of Design
Penn State University
16801, University Park, PA
scarlettmiller@psu.edu

Dean Parry
Pharmacies
Geisinger Health System
17822, Danville, PA
dparry@geisinger.edu

ABSTRACT

The US Institute of Medicine reports that one medication error occurs *per patient per day* in hospital care, and other studies indicate that medication administration errors attributed to packaging and/or labeling confusion can be as high as 33%. While many engineered products have identifiable features that help establish commonality and differentiation within a product family, vital features of consumable products such as medications are often not readily apparent in their physical form. As a result, caregivers must rely on the labeling and packaging to effectively determine the contents. Adverse Drug Events (ADEs) are the most common category of medical errors and include wrong drug, wrong dose, wrong route of administration, and wrong patient. It is estimated that in the US each year, medication errors harm at least 1.5 million people, resulting in 106,000 deaths. Computational models and associated decision support systems have the potential to improve pharmaceutical delivery safety through informed design of packaging features and enhanced situational awareness and decision-making during drug identification and administration. Past research has led to the formulation of measures for representing the degree of commonality and differentiation of packaging features in pharmaceutical families or versus look-alike drugs. Preliminary studies have validated these measures of feature prominence based on feature size and location. This paper describes a study using eye tracking to evaluate gaze patterns and further validate these measures. The results support the measures and indicate that increased commonality of features results in shorter reaction times, but also shorter fixation times. These results have implications in the formulation of a resulting decision support system.

Keywords

Pattern Recognition, Pharmaceutical Safety, Product Family Planning

1. INTRODUCTION

U.S. Pharmacia estimates that there are approximately 62.9 million medication dispensing mistakes a year in hospitals and pharmacies nationwide (Hicks, 2008), with just over three million of these mistakes considered to be medically significant. These significant mistakes can result from misreading labels, misprescribing, giving the wrong dose, or incomplete documentation. Classified as ADEs or "Adverse Drug Events," these events are defined and classified as "an injury due to medication management rather than the underlying condition of the patient" (Aspden, 2007). When in a hospital, it is estimated that a patient is on the receiving end of these mistakes an average of once per day, which can add up to $6,000 to a medical bill (Aspden, 2007). These errors can and do have detrimental effects to the patients' health and finances. Aside from the individual and familial effects that are seen from decreased health and death from illness or an ADE, there are large societal costs. These mistakes cost the nurses, the pharmacy and the health insurance providers. The accrued total of these mistakes must be distributed to cover the end cost to the patient as well as the healthcare companies. To do this, these costs are implemented throughout various industries, costing society as a whole not just the ones being directly

impacted by the mistakes. Often these mistakes occur because of the lack of consistency and regulation among package designs. Pharmacists and general consumers use cognitive decision-making processes to determine the correct medication, so by optimizing designs with this in mind, the number of mistakes may be able to be reduced greatly.

Ampuero and Vila performed a study on "Consumer Perceptions of Product Packaging" which focused on isolating different aspects of the labels on medical devices such as color, shape, image and typography (Ampuero and Vila, 2006). Isolating individual features on a label helped researchers begin to manipulate and understand how the brain is perceiving and processing the information on the label. Other aspects of the packaging of the medication itself can lead to a specific perception as to what volume is contained within it (Folkes and Matta, 2004). From a commercial manufacturer side, it is important to understand how the consumer views products on a shelf. Young (Young, 2012) presented the "PRS Eye Tracking Method" detailing how products on a shelf are viewed by customers. Understanding how commercial products are viewed and preferred is important specifically in package design, since consumers are generally novices in regards to looking at medication labels.

2. COMMONALITY/DIFFERENTIATION MEASURES

A powerful indicator of the importance of visual information can be its prominence as measured by size and location. To understand how the size and location can connote commonality, Shooter, et al. (2008, 2010) studied nine packages of Tylenol adult acetaminophen. The front panel (the side of the box that faces consumers when placed on store shelves) was analyzed in terms of its features, the text, and the graphics that conveyed information. The front panel was treated as a coordinate and normalized to account for variations in box size. Figure 1 shows the normalized face with each feature element identified and its centroid location. The area of each feature on the normalized package faces was calculated, tabulated, and compared relative to the entire area in order to reflect the package face "real estate" occupied by each feature as shown in Figure 2.

Presumably, features deemed by the manufacturer to be most important take up the most area on the package. The background uses the most space so one might expect consideration of color to be important. Tylenol uses a consistent shade of red for background color across its product family packaging. The brand name Tylenol takes up the second largest area, almost equal to the background. The type of medication and purpose are considerably smaller. Dosage and form are smaller still. The variation in normalized size and location of the Tylenol brand across the nine packages is minimal.

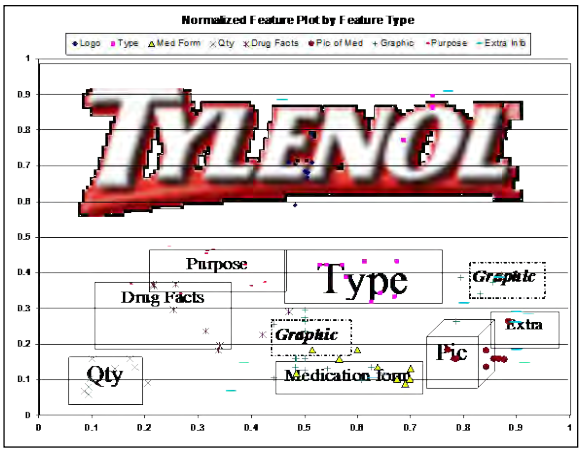

Figure 1: Front Panel Features' Positions

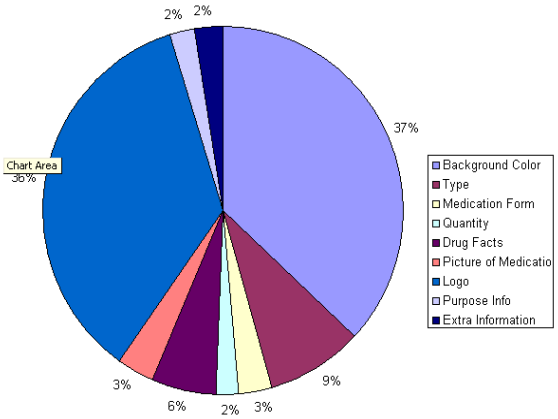

Figure 2: Relative Area of Front Features

There is also considerable consistency in location and size of the "secondary" information such as medication type, purpose, dosage, and form. It is evident that the Tylenol product family utilizes a package platform of features with an emphasis on brand recognition as the most salient common element. Information that differentiates among the variants of the product is less prominent in the package face, but is critical for informed and proper use. It is important to note that the Institute for Safe Medication Practice has recognized Tylenol as having a high number of cases of medication error (Hicks, 2008).

Cohen and Shooter (2010) followed this study with the formulation of measures to represent commonality and differentiation of packaging features with regard to prominence. In this preliminary investigation, the Feature Area Commonality Index (FACI) was formulated for packaging. One advantage of the commonality indices developed by Thevenot et al. (2007) and Alizon, et al. (2009) is that the result ranges between 0 and 1. The progression here is intended to provide a similar scale. The normalized areas for each

feature on each package was determined and tabulated. The mean was then calculated for the area of that feature across the package family. The *Proportion Difference* from the mean is calculated for each instance and represented as shown in Equation 1.

$$Proportion\ Diff = \frac{|\bar{A} - A_j|}{\left(\frac{\bar{A} + A_j}{2}\right)} \quad (1)$$

where $\bar{A}$ is the mean and A_j is the instance value

The measure can help identify and quantify an outlier instance in the packaging family. It is also possible to gain perspective on the total family by calculating the *Average Proportion Difference* for each feature using the value from each package variant. The intent is to describe the degree of commonality for features repeated across variants. If a feature is not present on a package instance, then that instance is not part of the calculation. A value of 1 indicates exact commonality across the variants while 0 means they are different. The *Feature Area Commonality Index (FACI)* is then calculated as seen in Equation 2 below.

$$FACI = 1 - Average\ Proportion\ Difference \quad (2)$$

The FACI provides a measure for each feature (i.e., Brand Logo) across the package family. The FACI for the Brand Logo feature was determined to be 0.92, indicating high commonality. The FACI for the Main Ingredient was 0.68 and for the Medication Use was 0.73, which are also strong indicators of commonality.

It can also be beneficial to represent the aggregate for all of the features. The most direct formulation of the *Aggregate Feature Area Commonality Index (AFACI)* is to take the average of all of the FACI values for the package family. In calculating the AFACI, only the feature categories present in the package family are included. For example, if the package family does not include any instances of the *Flavor Text* feature category, then all instances will have an area of zero, which results in a FACI of 1 (all commonly not present). If these absent features were included in the AFACI, then the result would be skewed toward a higher indication of commonality. This formulation considers all features as equivalent contributors; weighted contributions of different features will be investigated as part of this work. The AFACI for all features of the Tylenol package family studied was calculated to be 0.73, which is a strong indicator of commonality.

A similar approach was taken to formulate the Feature Location Commonality Index (FLCI), which is an indicator of the commonality and differentiation of the location of features across a product family based on clustered distances. We then validated both of these measures through a cognitive workload analysis study with 60 human subjects in Cho et al. (2014). Response time and selection accuracy were found to be positively correlated with the indices.

3. PREVIOUS VALIDATION STUDY

A Penn State University study, *Effects of Over-the-Counter Medication Product Family Design on Knowledge Acquisition and Consumer Preferences* (Cho et al, 2014) sought to study which features in Over the Counter (OTC) medical labeling specifically can be manipulated to encourage consumers to read and process the labels before they choose a medication for purchase. To do so, Robitussin and Equate labels were altered in Adobe Photoshop to create five different variations. These variations included a base design without any emphasized features, a design with increased font size, a label with inclusion of an accent color, a design with an addition of a graphical icon, and a variation with all of the emphasized features. The study was set up as a survey on Qualtrics software (Qualtrics, Provo, UT). Subjects were given symptoms and requested to select the corresponding medication. Accuracy and selection time were measured.

This study found that "variations in labeling and product family design significantly impacts the accuracy and efficiency of medication decision making and thus has the potential to reduce adverse drug events made during the process" (Cho et al, 2014). The study determined that the overall package design did not have a significant impact on the accuracy of a subject's selection. However, increased font size exhibited the shortest selection time, suggesting that increased font increases efficiency. The variation with all of the emphasized features had the longest selection time, which suggests that too many features could be distracting and decrease efficiency in selection. This could point to a limit in how many features can be emphasized before it detracts a feature's prominence and creates clutter. In looking at design recommendations in other avenues of academia, perhaps there is a "design magic number seven" that could be used in label design. Such a concept would limit how many features can be emphasized until they cancel out one another (Miller, 1956).

The study at Penn State also found that a higher commonality of both AFACI and AFLCI resulted in a higher accuracy of selections and a shorter selection time. Packages with this higher AFLCI, or higher commonality of locations of features across the product family, had a lower consumer preference rating from participants. The study concluded by noting that the task was more of a search task than a decision task. It suggested that a future study with an eye tracker or employee tasks more closely related to the decision making process of selecting medications would be appropriate direction to pursue in the future.

Additionally, it was suggested that future studies test a wider range of OTC labels to create a normal distribution for the indices and use product-moment correlation.

This paper investigates and expands upon Cho et al, performed in 2014. It examines areas of particular interest and importance to individuals interacting and giving medications. Within product families, commonality of the specific features of interest were calculated and then compared against accuracy and viewing patterns to further understand how information is processed from medical labels. The study validates the efficacy of the commonality and differentiation measures for visual features on pharmaceutical packaging. Pattern recognition techniques and automated calculation of these measures will support the rapid exploration of alternative designs with the goal of improved dispensing of medications.

4. EXPERIMENTAL PROCEDURE

Participants

Fifty-four undergraduates from a small college in Pennsylvania were used as the participants of this study. Half of the participants were assigned and exposed to the Robotussin and the other half to the Equate stimuli pool. They were part of a psychology class subject pool and received credit for participation.

Materials and procedure

Stimuli used in the Cho study (Cho et al., 2014) were recreated using Adobe InDesign (Viers et al, n.d.) in the same configuration and patterns. The types of labels used included base (with no manipulations), coloring, icon of person, font size and an image with all available manipulations. A total of 50 images were created that each consisted of 4 vertical labels placed side by side with a corresponding question running along the top in the white space. The questions were also recreated and matched from the previous study (Figure 3). The images were then placed into TOBI software (TOBI, Fall Church, VA), 25 Robotussin images and 25 Equate images in their own respective trial setups.

The images that were uploaded into the program were then processed with areas of interest being identified and outlined using the TOBI software (TOBI, Fall Church, VA). These Areas of Interest (AOI) were as follows: brand name, comparison, name, moon/nighttime (specifically used for Robotussin), description, active ingredients, non-drowsy, symptoms, button, and dots/indicators on the body that the drug will apply to. The labels are coded in the following manner: "ImageID#.AOI#.correct/incorrect". The labels were each individually coded as "correct" or "incorrect" relative to the answer to the question to be able to separate the data as such. A sample marked up image can be seen in Figure 4.

Figure 3: Sample stimuli with question of equate, used in the study.

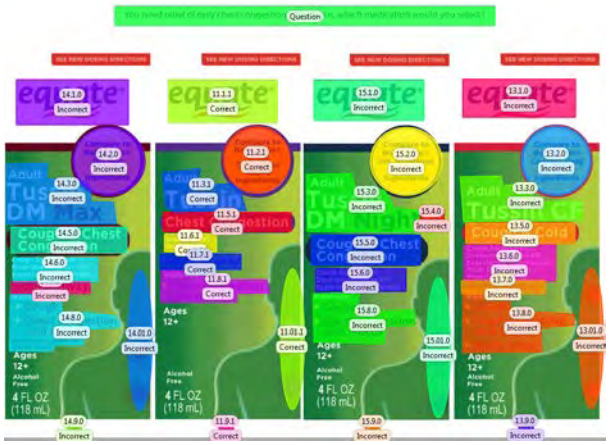

Figure 4: A sample marked up image of a quartet.

Upon entering the study, subjects were given the "Functional Health Literacy in Adults" survey as well as a few general questions about their educational background. This was used to measure how much effort and attention the subject were giving during the study. Afterwards, the TOBI software (TOBI, Fall Church, VA) was opened and the eye-tracking tool was calibrated to the subject. Once completed, subjects were instructed that they were going to be shown images and they would have to read the questions and click on the "correct" image by selecting the circles underneath the corresponding label. Once the subjects had seen all 25 images in their assigned trial, the study was concluded. A sample viewing pattern displayed via heat map can be seen in Figure 5.

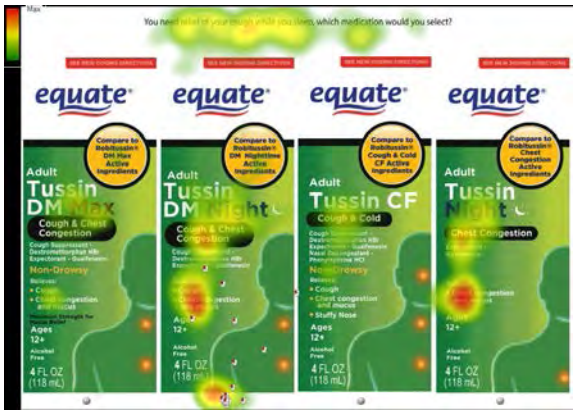

Figure 5: A heat map of one subject's viewing of an image displayed during the study. Yellow indicates shorter viewing times, red displays a longer view time as indicated in the key in the top left corner of the image.

Images were analyzed using the method utilized by (Cho et al, 2014) calculating both location and area commonality of each feature on the label and comparing them within their product families. With this method, product family metrics were used when analyzing the correlation between package design and selection results. Feature Area Commonality Index (FACI) and Feature Location Commonality Index (FLCI) were used to describe commonality of an area of interest. Feature Area Commonality Index (FACI) was determined by calculating the physical area of a label of each specific feature covered.

The FACI was then calculated by subtracting the result from equation 1 from 1, mentioned previously in equation 2. This is useful as if the commonality is consistent across the product family, the FACI is close to 1, and if they are completely differentiated, the result is 0.

Feature Location Commonality Index (FLCI) was determined in a similar manner to the FACI, but the location was determined from the left edge of the package to the centroid of the feature. The locations were again averaged across product families and then subtracted from 1 to calculate the FLCI (Cho et al, 2014). This specific task allowed for analysis of the commonality of location across groups. Each of these values, FACI and FLCI, calculated for feature across product families were then averaged to create aggregates of the FACI and FLCI, further referenced as the AFACI and AFLCI. The AFACI and AFLCI give a good indication as to the commonality across the product family.

Results

An item analysis was conducted, rather than analyzing by subject, so that the effect of commonality on gaze patterns for individual label features (e.g. brand name, active ingredients) could be investigated. For each image, the following were analyzed: the time to first fixation, fixation count, and mean fixation duration for each area of interest (AOI; see Method for description). AOIs that received no fixations were not included in the analysis. The area and location commonality indices (FACI & FLCI, respectively) reported in Cho et al. (2014) for each AOI (see Table 1) were also compared.

A multivariate analysis of variance (MANOVA) was conducted on these five variables across the AOI regions. Results of the MANOVA reveal a significant multivariate effect, Pillai's Trace=1.28, F (30, 1050)=12.08, p<.001, η_p^2=.257, demonstrating a difference in the commonality and pattern of eye movements for the regions of interest. Univariate analyses for this relationship indicate a significant effect of AOI on time to first fixation (F(6,210)=73.36, p<.001, η_p^2=.68), mean fixation duration (F(6,210)=7.78, p<.001, η_p^2=.18), fixation count (F(6,210)=48.33, p<.001, η_p^2=.58), FACI (F(6,210)=15.47, p<.001, η_p^2=.31), and FLCI (F(6,210)=28.81, p<.001, η_p^2=.45).

The primary analysis of interest was the relationship between each feature's commonality (within a product family) and eye-gaze patterns; thus, we conducted a series of bivariate correlations (6 comparisons, Bonferroni corrected α=.008). Both FACI and FLCI were significantly negatively correlated with time to first fixation (FACI: r(216)= -.418, p<.001; FLCI: r(262)= -.535, p<.001; see Figure 1), indicating that for AOIs with greater commonality, participants fixated on that AOI earlier in the trial. In addition, FACI and FLCI were significantly negatively correlated with fixation count (FACI: r(216)= -.354, p<.001; FLCI: r(262)= -.450, p<.001; see Figure 7) and mean fixation duration (FACI: r(216)= -.305, p<.001; FLCI: r(262)= -.193, p=.002; see Figure 8). This suggests that features with greater commonality received fewer fixations and those fixations were shorter in duration.

Table 1. Means and standard deviations of the eye-tracking metrics and commonality indices for the 7 prominent AOIs.

	Time to First Fixation (ms)		Fixation Count		Fixation Duration (s)		AFACI		AFLCI	
Feature/AOI	M	SD	M	SD	M	SD	M	SD	M	SD
Brand Name	29.52	23.64	1.65	1.64	0.31	0.34	0.9793	0.01	0.9907	0.01
Name	94.59	24.98	7.40	3.47	1.13	0.39	0.7067	0.11	0.9391	0.01
Description	102.70	35.66	6.79	2.32	1.00	0.26	0.7987	0.19	0.8856	0.00
Active Ingredients	70.47	37.36	2.87	1.62	0.55	0.41	0.6938	0.26	0.9415	0.04
Non-Drowsy	71.52	19.50	2.28	0.79	0.39	0.15	0.9061	0.12	0.9229	0.01
Symptoms	144.42	32.52	14.73	4.93	1.33	0.45	0.6232	0.32	0.9027	0.06
Dots	16.81	9.88	0.43	0.24	0.10	0.05	0.9686	0.02	0.9681	0.02

Note: Table 1 only displays AOIs that had commonality indices reported in Cho et al. (2014).

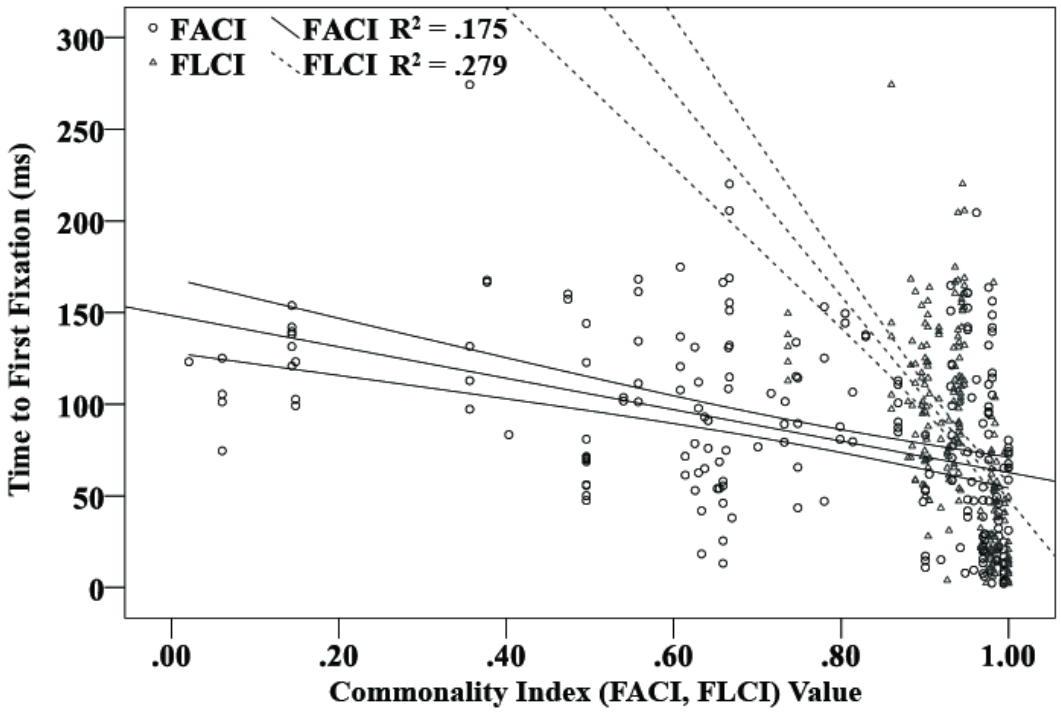

Figure 6. Scatterplot showing relationship between commonality (FACI and FLCI) and time to first fixation (ms).

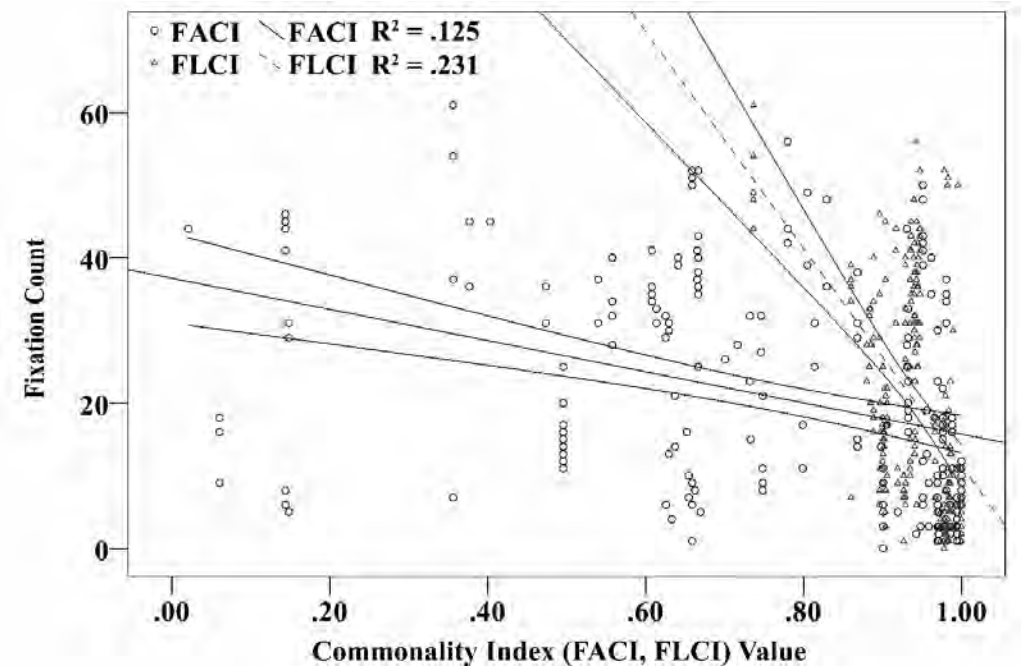

Figure 7. Scatterplot showing relationship between commonality (FACI and FLCI) and mean fixation count across areas of interest (AOIs).

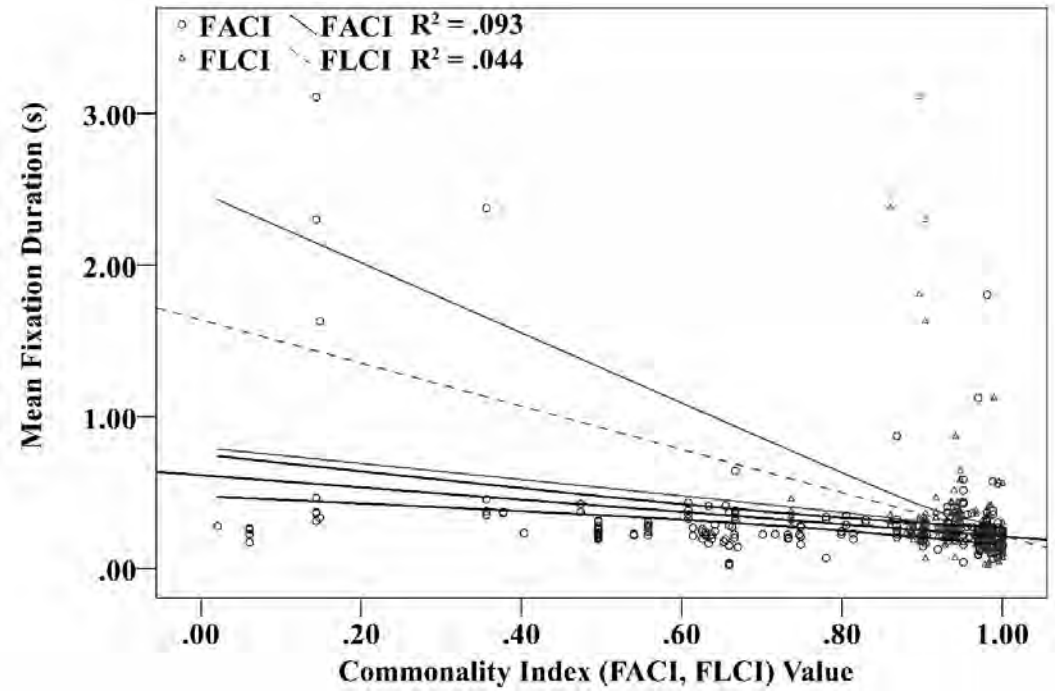

Figure 8. Scatterplot showing relationship between commonality (FACI, FLCI) and mean fixation duration across areas of interest.

5. DISCUSSION AND IMPLICATIONS

Examining the results, it is supported that as subjects become more familiar with the location of certain pieces of information, or specific features on the labels, less time is spent looking at these features. The significance of the correlations we found between the FLCI and FACI reflect this hypothesis. From this data, it is suggested that increasing commonality values among product families, in both location and area, will help to decrease reaction time. An increase in commonality could be applied to both over the counter medication as well as pharmaceuticals.

A shorter reaction time could be helpful and harmful. Decreased reaction time may allow for the pharmacist or consumer to attain more information from the label in a shorter amount of time, or it could decrease the saliency of the information as less time is being spent reading the label. More information in less time would ideally mean fewer mistakes would be made, as the majority of the information on the label would cross through the handler's field of vision. However, more information could also mean an overload of information in the field of vision so less of it is retained. Commonality and regulation among product families affect nurses administering drugs in the hospital setting. For example, if the same information (such as dosage) was in a particular location on every vial the nurse would have to exert minimal effort to determine what the dosage was any given vial, but perhaps they would then minimally view the numerical information given by the feature.

With the knowledge that commonality helps to increase reaction times, future studies will to take this data and build on this concept. A similar study will be performed using prescription labels and medical staff. Since the previous subject pool uses novice medical label viewers, this study would ideally help to support the idea that commonality influence novices as well as experts with years of training. This subject pool is exposed to medical packaging every day, and would be most likely to notice or be affected by changes in labels.

Moving forward new labels will be created, with information that has been manipulated to contain high or low levels of commonalities. The different information found in these areas will test specifically if high levels of commonalities decrease salience of information. It is important to see how the location of information influences the accuracy and consistency of information retrieval as opposed to just reaction time. Results of the current study support that there is less time spent in areas with high commonality, but the researchers are currently unable to determine how salient the information is in those high commonality locations. If commonality decreases reaction time, than perhaps the information provided in these high commonality areas are not as salient as originally intended.

To improve validity of results, comparing two-dimensional renderings of labels and a more realistic rendering is of interest to determine which format will provide a more accurate replica of what nurses would experience in the field handling medical vials. Future studies will investigate the brightness of a computer screen versus brightness of physical labels and how the lighting of testing scenarios might affect the ability for results to extend to a physical pharmacy or hospital dispensary. Studies could also delve into color theory and examine how lumosity might influence choices of medical labels. With a commonality in lumosity or actual color within the feature, perhaps saliency would be able to increase with commonality.

6. IMPLICATIONS FOR DECISION SUPPORT SYSTEM

The study has validated the saliency of the measures for commonality and differentiation of packaging features with improved medication selection. There are tens of thousands of medications on the market both over-the-counter and prescription. The intent is to create a database of diverse pharmaceutical packages that will include the Packaging Commonality Differentiation Indices and highlights of identified "trouble spots". This information will be used to improve upon the current pharmacy dispensary approach where red labels are used as warnings for potential identification hazards, as well as to improve on the internal labeling used for in-patients. The computational models for creating the indices in the database need to be automated as much as possible to relieve burden on entry of new package information. For example, the information capture has been simplified, and the computation of the FACI and FLCI indices has been automated by using software that automatically measures the area and centroid locations of features; and then calculates the indices. However, it is desired to automate the recognition and categorization of features supplied to the measures. Techniques of pattern recognition and cluster analysis will dramatically improve this process. The intent is to develop a system that will provide information to a package designer that will enable the rapid exploration of alterative designs with improved medication administration outcomes.

7. CONCLUSIONS

This paper has introduced a computational model for representing the commonality and differentiation of visual features on pharmaceutical packages. The measures had been previously validated through a

workflow analysis study. This study used eye tracking to evaluate gaze patterns for novice subjects. The results support the measures and indicate that increased commonality of features results in shorter reaction times, but also shorter fixation times. A similar study is currently being conducted with healthcare professionals including nurses and pharmacists to explore the correlations in healthcare settings. The intent is to develop a working decision support system that will support the exploration of alternatives to packaging designers. The implications for decision support in organization is also being examined as well as the structure of pharmaceutical dispensing. The researchers have taken the approach of validating the measures before the development of the decision support system to ensure the efficacy of the approach. The calculation and representation of the measures have been automated. The team is currently exploring techniques for pattern recognition to automate the recognition and categorization of salient features.

8. ACKNOWLEDGEMENTS

This research is supported by the Bucknell Geisinger Research Initiative.

9. REFERENCES

Alizon, F., Shooter, S. B., and Simpson, T. W., 2009, "Assessing and improving commonality and diversity within a product family," Res. Eng. Des., **20**(4), pp. 241–253.

Ampuero, O., and Vila, N., 2006. "Consumer perceptions of product packaging," J. Consum. Mark., 23(2), pp. 100-112.

Aspden, P., Wolcott, J., Bootman, J. L., and Cronenwett, L. R., 2007, Preventing medication errors, National Academies Press Washington.

Cho. J., Miller, S., Simpson, T., and Shooter, S., 2014, "Effects of over-the-counter medication product family design on knowledge and acquisition and consumer preferences," Proceedings of the ASME 2014 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference (IDETC/CIE), Buffalo, NY.

Folkes, V., and Matta, S., 2004. "The Effect of Packaging Shape on Consumers' Judgements of Product Volume: Attention as a Mental Contaminant," J. Consum. Res., 31(2), pp. 390-401.

Hicks, R. W., Becker, S. C., and Cousins, D. D., 2008, ""MEDMARX data report: report on the relationship of drug names and medication errors in response to the Institute of Medicine's call for action,"" Rockville, MD: Center for the Advancement of Patient Safety, US Pharmacopeia.

Miller, G., 1956. The Magical Number Seven. The Psychological Review.

Qualtrics, (2015). Qualtrics. Provo, UT.

Shooter, S. B., Cohen, S., and Williams, C., 2008, "Assessing Commonality and Differentiation for Packaging Family Planning with Application to Medication Labels," Proceedings of the ASME2008 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference (IDETC/CIE), pp. 67–76.

Shooter, S. B., and Cohen, S. W., 2010, "The Commonality Differentiation Index Using Prominence of Visual Information for Medication Package Family Planning," Proceedings of the ASME 2010 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference (IDETC/CIE), Montreal, Quebec, Canada, pp. 231–237.

Thevenot, H. J., Alizon, F., Simpson, T. W., and Shooter, S. B., 2007, "An Index-based Method to Manage the Tradeoff between Diversity and Commonality during Product Family Design," Concurr. Eng., **15**(2), pp. 127–139.

Viers, R., Ramos, J., Koren, G., & Brodnitz, D. (n.d.). Adobe InDesign CS5 [Computer software].

Young, S., 2012, "Rigid Plastic: Applying an Architecture," Packag. Des. Mag.