Package Design Affects Accuracy Recognition for Medications

Tor Endestad, **Laura A. Wortinger**, University of Oslo, Norway, **Steinar Madsen**, and **Sigurd Hortemo**, Norwegian Medicines Agency, Oslo, Norway

Objective: Our aim was to test if highlighting and placement of substance name on medication package have the potential to reduce patient errors.

Background: An unintentional overdose of medication is a large health issue that might be linked to medication package design. In two experiments, placement, background color, and the active ingredient of generic medication packages were manipulated according to best human factors guidelines to reduce causes of labeling-related patient errors.

Method: In two experiments, we compared the original packaging with packages where we varied placement of the name, dose, and background of the active ingredient. Age-relevant differences and the effect of color on medication recognition error were tested. In Experiment 1, 59 volunteers (30 elderly and 29 young students), participated. In Experiment 2, 25 volunteers participated.

Results: The most common error was the inability to identify that two different packages contained the same active ingredient (young, 41%, and elderly, 68%). This kind of error decreased with the redesigned packages (young, 8%, and elderly, 16%). Confusion errors related to color design were reduced by two thirds in the redesigned packages compared with original generic medications.

Conclusion: Prominent placement of substance name and dose with a band of high-contrast color support recognition of the active substance in medications.

Application: A simple modification including highlighting and placing the name of the active ingredient in the upper right-hand corner of the package helps users realize that two different packages can contain the same active substance, thus reducing the risk of inadvertent medication overdose.

Keywords: human error, package labeling, designing for the elderly, memory, cognition, patient safety

INTRODUCTION

Medication error is a major patient safety issue in the United States with 1.5 million adverse drug events reported annually, over one third of which occur in the outpatient setting, at an annual estimated cost approaching $1 billion. Over-the-counter (OTC; i.e., nonprescription) drug use is increasing; almost one half of U.S. adults take at least one OTC medication regularly. Almost one fifth of U.S. adults take acetaminophen in any given week (Wolf et al., 2012). Acetaminophen overdose is the leading cause of liver failure in the United States, and the package labeling of acetaminophen-containing OTC medications is a likely contributor to many unintentional overdoses (Wolf et al., 2007, 2012). A substantial number of medication errors may be related to name confusion due to inadequate labeling on medication packaging. Labeling effectiveness may be influenced by placement of the drug name and dosage strength, undue prominence of company logos compared with drug identifiers, or similar-looking labels or packaging design attributes on different products (Berman, 2004; Estock et al., 2015; Kongkaew, Noyce, & Ashcroft, 2008).

In a recent review, Mira, Lorenzo, Guilabert, Navarro, and Pérez-Jover (2015) found that medication error rates were between 19% and 59% in studies from different countries. Even if many of the errors do not lead to life-threatening situations, such errors are the principle cause of adverse events, especially among the elderly (Weiss, 2009). This user group also elicited the largest number of errors (Fraeyman et al., 2015; Mira, Lorenzo, et al., 2015; Weiss, 2009). The elderly population is particularly at risk for consequences of medication errors (Estock et al., 2015; Fraeyman et al., 2015; Weiss, 2009), because it both uses more medications and is more vulnerable to adverse drug events. According to Holden, Schubert, and Mickelson (2015),
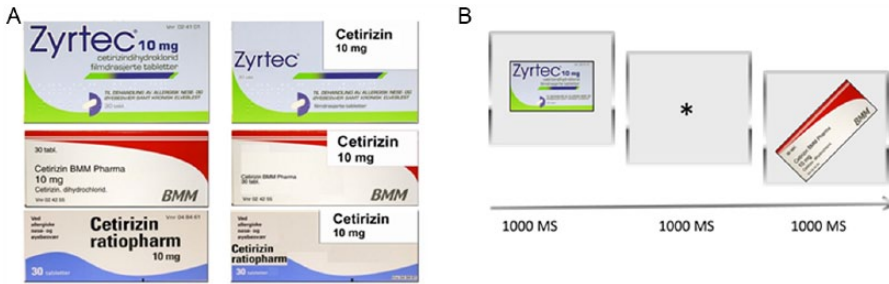
Figure 1. (A) Example of original (left) and redesigned (right) packages. (B) A trial consisted of a reference image, a fixation cross, and the target image. By pressing a key, participants indicated whether the target image contained the same or a different active ingredient as the reference image.

in the United States, 80% of older adults have at least one chronic disease and 50% have two or more, accounting for 75% of health care expenditures. Because elderly users of medication are increasing as the population ages, they are particularly at risk for making medication-related errors.

Of particular concern is the effect of increased use of generic drugs on adherence (Håkonsen, Eilertsen, Borge, & Toverud, 2009; Kjoenniksen, Lindbaek, & Granas, 2006). In an interview study of 174 Norwegian hypertensive patients, Håkonsen et al. (2009) found that 5% of the patients used more than one equivalent generic product at the same time. One of the factors identified by the patients as leading to these errors was their belief that the generic drug was an additional drug because they thought the name of the manufacturer was the name of the drug. In a follow-up study, Håkonsen and Toverud (2010) found that 10% of a sampled Pakistani population in Oslo, Norway, used two equivalent generic drugs at the same time after the introduction of generic substitution in pharmacies. These findings have been replicated in a study of unintentional overdose with acetaminophen products (Wolf et al., 2012), whereby 24% of adult users demonstrated that they would overdose and 46% would take a "double dose" of two identical generic products.

Existing recommendations for the optimal design of medication labels are based on expertise and sound human-centered design principles. These recommendations include how and where to print the drug name, placement and prominence of company logos, and use of tall letters to support

name discrimination (Bailey, Pandit, Curtis, & Wolf, 2009; Berman, 2004; Chafac & Chan, 2012; European Parliament & Council of the European Union, 2001; Filik, Purdy, Gale, & Gerrett, 2006; Fraeyman et al., 2015; Helen Hamlyn Research Center & National Safety Agency, 2007; Kenagy, & Stein, 2001; Ward, Buckle, & Clarkson, 2010; Wolf et al., 2012; Yin et al., 2010). However, as Estock et al. (2015) state, there is little empirical evidence to support these recommendations. We set out to provide empirical support for the use of substance name and dose as the main source of information and for more consistent placement of this information. On our redesigned packaging, this information was placed in the upper right corner of the packages (see Figure 1B for examples) to avoid confusion with brand names and other proprietary information and in line with human factors studies using eye tracking to identify optimal information placement (Arnheim, 2004; Rajashekar, Cormack, & Bovik, 2002).

To compare the redesigned packages with the original packages, we wanted to provide a task that took into account both mental effort and the fact that medication packages might be viewed from different angles relative to the users. We chose a modified version of the Shepard and Metzler (1971) mental rotation task (MRT). The MRT has been used to evaluate a range of design issues in human-centered design (Aitsiselmi & Holliman, 2009; Barfield, Sandford, & Foley, 1988; Cooper & Podgorny, 1976; Hancock, Carpendale, & Cockburn, 1988). The MRT is a well-established example of a recognition working-memory task that allows manipulation

of mental workload (Barfield et al., 1988). Because users may have varying degrees of cognitive resources available when trying to recognize medication packaging, the MRT appears to be an effective way to evaluate medication package recognition. In addition, given that medication packages are usually either held or placed at different angles, the visual frame through which they are evaluated will change similar to the MRT task.

Since color can be a source of confusion (Brandt, 2015; Cohen, 2000), we wanted to test the robustness of the proposed design principles in situations with confusing color schemes. We sought to test the robustness of the suggested design for errors related to color confusion (i.e., the same color scheme used on different medications, leading users to believe they are in fact the same substance). We therefore included a color factor with two levels: same- and different-colored packages.

We also wanted to test the robustness of the design with regard to age-related challenges. Age is associated with decreased capabilities that are important for identifying and recognizing the correct medication, for instance, reduced working-memory performance, attention, speed of information processing, and inhibitory control (Glisky, 2007). In addition, both vision and perceptual function in general decline with age (Baltes & Lindenberger, 1997) and can interact with attentional resources in such a way that elderly users might perform worse than might be expected from reduced cognitive function alone. For recognition tasks, age is related to lower visual accuracy and longer reaction times (Mira, Guilabert, et al., 2015). Because it is important to know whether design changes that reduce errors for young users may also do so for elderly users, we included a group of users older than 70 years.

We studied three drug package manipulations—standardized placement of information, light background, and highlighted substance name—in a task requiring a decision between brand- and generic-named medications. On the basis of the literature, we predict that the proposed design will lead to faster and more accurate identification of medications in both elderly and young users. Second, we predict that both the use of substance name and its

placement will affect discrimination. Finally, we predict that the redesigned packages will be less prone to being misread due to age and color confusion.

EXPERIMENT 1

Method

Participants. Fifty-nine volunteers (38 female) within two age groups were recruited and compensated to participate in this study, including 30 elderly users (69–86 years, mean 75.9; 20 females) and 29 young students (18–38 years, mean 25.9; 18 females). Each gave written informed consent in accordance with protocols approved by the University of Oslo, Norway. The participants were interviewed regarding their use of substances. The participants in the elderly group used on average 3.1 drugs, whereas the students used 0.5 different drugs. This research complied with the tenets of the Declaration of Helsinki and was approved by the Institutional Review Board at Department of Psychology, University of Oslo.

Materials. Images of 20 pharmaceutical packages with the original design and 20 packages with a manipulated design were created with a Canon camera (see example in Figure 1). The redesigned packages were made using Adobe Photoshop CS2. We inserted a white rectangle (occupying 25% of the front) in the upper right-hand corner of the package. Here we displayed the name of the active ingredient (Arial bold 24-point font size) and dose (using Arial bold 18-point font size). On some small packages, the fonts were reduced proportionally. The brand names were reduced to 60% of original size.

Design. Participants were presented with an MRT, a classic cognitive test used to measure recognition and discrimination performance in individuals. We used an adapted version of the original MRT task described by Shepard and Metzler in 1971 (Peters & Battista, 2007; Pinker & Finke, 1980; Shepard & Cooper, 1982; Shepard & Metzler, 1971).

We used packages in two versions, one with original names, labels, and colors (original) and one with our redesigned information placement (redesigned). In addition, we created versions of

the packages with new colors to enable a factorial measure of the same substance with same (non-interfering) or different (interfering) color as well as different substance with same (interfering) and different (non-interfering) color.

Procedure. The elderly participants were recruited from senior citizens' day centers. The testing was done in a laboratory setting. The participants signed a consent form and were shown the packages on laptop computer screens positioned at a comfortable distance in front of participants. They were shown examples of the task and did a brief training session before they completed the 720 trials divided into two sessions.

In each trial, we presented a Picture A (reference image) for 1 s, followed by a fixation cross for 1 s, and finally, a Picture B (target image) for 1 s (Figure 1). The participants were allowed to reply within an intertrial interval of 3 s.

The pictures were viewed from about 70 cm on a white background. The drug packages were presented in the center of the display. The second picture was presented at rotations of 0°, 60°, 120°, or 180°.

In each trial, the participants had to indicate as quickly as possible whether the two packages contained the same active ingredient. By pressing a green (same) or red (different) key on the computer keyboard, participants placed each index finger on one key and had a reaction time cutoff of 3 s.

Data analysis. Data analysis was performed with SPSS v19. The experimental conditions age (young, elderly), design (original, redesigned), and substance (same, different) were applied in a mixed within- and between-group analysis. To correct for multiple comparisons, Bonferroni-corrected post hoc analysis with a cutoff of .05 was used.

Only reaction times for correct responses were included in the analysis. Because of known age-related effects on reaction time performance of MRTs (Dror, Schmitz-Williams, & Smith, 2007), we calculated slope and intercept of reaction times to avoid confounds. Although there was no effect of age on reaction time slope, there was a significant difference in intercept (elderly, $M = 1,053$ ms; young, $M = 740$), $F(1, 57) = 43.966, p < .001, \eta^2 = .440$. Intercept correlated moderately with the mean reaction times (r^2 ranging from .41 to .55) and was included in the analysis as a covariate in the reaction time analysis.

Results

Reaction times. We performed a 3 × 2 repeated-measures analysis of reaction times using the intercept of the MRT as covariate. A significant main effect of age, $F(1, 57) = 54.594, p < .001, \eta^2 = .494$; design, $F(1, 57) = 75.834, p < .001, \eta^2 = .575$; and substance, $F(1, 57) = 46.725, p < .001, \eta^2 = .455$, was found. In addition, a significant two-way interaction was found between design and substance, $F(1, 57) = 13.247, p < .001, \eta^2 = .191$ (see Table 1). These interaction effects were due to a significantly larger difference between redesigned and original packages when the substances were the same compared with packages with different substances.

Accuracy. To compare the redesigned and original packages, we performed a two-way repeated MANOVA with design, substance, and age group as factors. We found a main effect of design, $F(1, 57) = 223.659, p < .001, \eta^2 = .795$; an interaction between design and substance, $F(1, 57) = 171.205, p < .001, \eta^2 = .750$; a main effect of age, $F(1, 57) = 42.094, p < .001, \eta^2 = .425$; and an interaction between design, substance, and age, $F(1, 57) = 4.988, p = .029, \eta^2 = .080$ (see Figure 2). These effects were due to a significant difference between redesigned and original packages when substance was the same (see Table 1 for confidence intervals for all the conditions) both for elderly and young subjects.

There was no significant difference between the groups for redesigned packages in this condition.

As discussed previously, a prominent contributor to adverse drug events is when a patient erroneously takes a generic drug containing the same substance as one he or she already takes ("double dipping"). To test the extent of the double-dipping error, we analyzed the errors in the condition where the packages contained the same substance as the reference sample. This method allowed for comparison of identical packages versus packages with the same substance but different brand.

In separate 2 × 2 MANOVAs, a significant two-way interaction between design and brand

TABLE 1: Mean Reaction Times, Accuracy, Standard Errors, and Confidence Intervals (CI) for the Different Versions of Packages

	Young		Elderly		
Variable	M (SE)	95% CI	M Correct (SE)	95% CI	p
Reaction time in milliseconds					
Original same substance	981 (37)	[907, 1055]	1282 (35)	[1212, 1352]	***
Original different substance	1036 (42)	[951, 1121]	1344 (36)	[1273, 1416]	***
Redesigned same substance	779 (33)	[713, 844]	1152 (31)	[1090, 1214]	***
Redesigned different substance	894 (38)	[819, 970]	1318 (40)	[1237, 1398]	***
Accuracy in percentages					
Original same substance	65 (4)	[56, 73]	44 (3)	[38, 50]	***
Original different substance	95 (1)	[94, 96]	77 (4)	[69, 84]	***
Redesigned same substance	92 (2)	[89, 96]	85 (2)	[81, 90]	**
Redesigned different substance	95 (1)	[93, 96]	78 (3)	[71, 85]	***

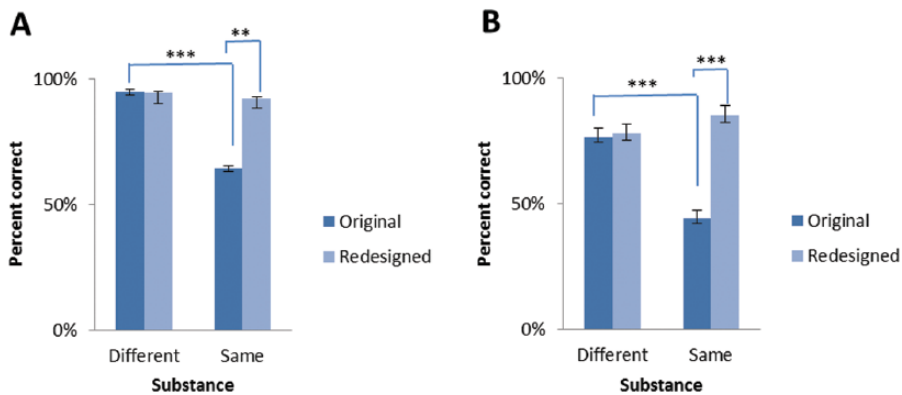$*p \leq .05.$ $**p \leq .01.$ $***p \leq .001.$

Figure 2. (A) Overall percentage correct responses for the two designs for the young group. (B) Overall percentage correct responses for the two designs for the elderly group.

was found in both age groups: elderly, $F(1, 29) = 110.104$, $p < .001$, $\eta^2 = .792$; young, $F(1, 27) = 27.000$, $p < .001$, $\eta^2 = .691$. The effect was due to a significant difference between redesigned and original packages when packages were generic and non-identical brands.

For the packages with different substances, only an effect of age was found, $F(1, 57) = 21.784$, $p < .001$, $\eta^2 = .276$, with elderly users making more errors than younger users.

Color. There are two conditions in which color can create confusion. The first is when two packages contain the same substance but have different color schemes. This design may lead users to the

erroneous conclusion that the packages contain two different substances. The second is when two packages contain different substances but the packages have a similar color scheme, leading the users to the erroneously believe that the packages contain the same substance. We tested whether the redesigned packages made such color-induced errors more or less likely when comparing different generic medications. A Package Design × Color × Age repeated-measures analysis for both color contexts showed that in the different-substance condition, a significant effect of age was found, $F(1, 57) = 20.611$, $p < .001$, $\eta^2 = .266$ (Table 2), but no significant effect of design.

TABLE 2: Error Rates in Percentages for the Original and Redesigned Packages When Users Matched Different Substances With Same or Different Color Schemes in Generic Medications

	Young		Elderly		
Variable	M (SE)	95% CI	M (SE)	95% CI	p
Different color					
Redesigned	6 (1)	[4, 8]	24 (4)	[16, 32]	***
Original	5 (1)	[4, 7]	25 (4)	[17, 33]	***
Same color					
Redesigned	9 (2)	[5, 13]	19 (2)	[14, 24]	**
Original	44 (5)	[34, 54]	69 (4)	[61, 76]	*

Note. CI = confidence interval.
*$p \leq .05$. **$p \leq .01$. ***$p \leq .001$.

For the same-substance and different-color conditions, there was a main effect of age, $F(1, 57) = 20.400$, $p < .001$, $\eta^2 = .264$; and design, $F(1, 57) = 206.220$, $p < .001$, $\eta^2 = .783$; and a significant interaction between age and design, $F(1, 57) = 6.183$, $p = .016$, $\eta^2 = .098$ (Table 2). The interaction effect was due to a significantly greater difference between redesigned and original packages for the elderly group.

The results point to color as an important source of confusion when two packages contain the same substance but are colored differently. Especially, elderly users erroneously conclude that the packages contain two different substances, which suggests that color might contribute to the "double-dipping" problem. Even though the redesigned packages seem to be more robust, elderly users made significantly more errors related to color than did the young users.

Discussion

Our main findings indicate that there are advantages related to both effort (reaction times) and accuracy for the redesigned packages in the MRT. However, since we highlighted the substance name with a prominent font, provided a distinct contrast background, and standardized placement of this information in Experiment 1, we cannot discern the role of these three manipulations. We therefore designed a follow-up experiment whereby we manipulated placement and contrast background for the same subset of packages as in Experiment 1. The main findings in Experiment 1

pointed to main effects of age and age interactions related to decrease in performance. The elderly performed worse in all the conditions. However, there were no interactions indicating a different pattern of errors in the two groups. We therefore chose to do the second experiment only with a group of young users.

EXPERIMENT 2

Participants. Twenty-five volunteers (13 female), all young students (21–28 years, mean 25.9), participated in the study. Each gave written consent in accordance with protocols approved by the University of Oslo, Norway. The participants were interviewed regarding their use of substances (0.6 different drugs).

Materials. The same materials were used as in Experiment 1. In the package-redesigned condition, the packages were identical to those used in Experiment 1 (see Figure 3). In the placement condition, we constructed a version of the redesigned packages with the white field and substance names placed in the lower left corner. In the third condition, the transparent condition, we prepared another set of packages with identical substance names positioned in the standard place (upper right corner) but with the original (not redesigned) color background.

Design. The Experiment 1 task was used. Experiment 2 had three conditions. In the redesigned condition, we used the same packages as in Experiment 1. In the transparent condition, substance names were placed in the upper right

TABLE 3: Mean Reaction Times in Milliseconds, Errors in Percentages, Standard Errors, and Confidence Intervals (CI) for the Different Conditions

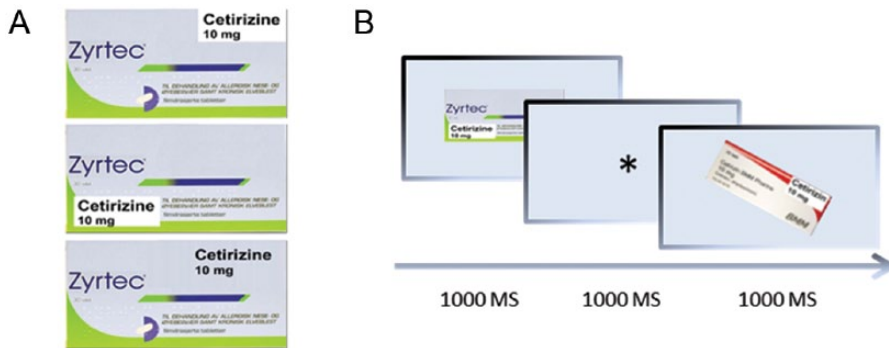| Condition	Reaction Time		Errors	
	M (*SE*)	95% CI	% (*SE*)	95% CI
Redesigned	1017 (40)	[934, 1100]	6 (1)	[4, 8]
Substance	1131 (47)	[1035, 1227]	12 (2)	[8, 15]
Placement	1035 (50)	[932, 1138]	7 (1)	[5, 9]

Figure 3. (A) Example of the three conditions: redesigned (top), placement (middle), and transparent (bottom). (B) The same trial procedure was used as in Experiment 1.

corner but without the contrasted background. In the placement condition, we used the same design as in the redesigned condition but randomly placed the critical information in either the upper right or the lower left corner of the packages.

Procedure. The procedure was identical to Experiment 1.

Analysis. The same statistical procedures were followed as in Experiment 1.

Results

Reaction times. A repeated-measures MANOVA showed a significant main effect, $F(2, 48) = 15.803$, $p < .001$, $\eta^2 = .397$ (see Table 3). The effects were due to a significantly longer reaction time for the transparent condition compared with the others. There were no differences between the redesigned and placement conditions.

Accuracy. A repeated-measures MANOVA showed a significant main effect, $F(2, 48) = 5.925$, $p = .005$, $\eta^2 = .198$ (see Table 3). The effects were due to more errors for the transparent condition compared with the other two conditions. The rede-

signed and placement conditions were not significantly different ($p = .590$).

DISCUSSION

There was a high rate of recognition error for generic medication in their original package designs. The most prominent error was to wrongly assume that the medication packages contained different substances when in fact they contained the same substance (elderly, 55%, and young, 35%). In addition, this type of error increased to 68% (elderly) and 43% (young) when the packages had different colors. Our experiment was done in an artificial experimental context and cannot predict the incidence of these types of medication errors in the real world. However, our results are consistent with documented field trials (Wolf et al., 2012). In a review of articles published between 1990 and 2014, Mira, Lorenzo, et al. (2015) found that the incidence of medication errors was between 19% and 59% and that the elderly users made more errors than young users. Incorrect dosage,

forgetting or mixing up medications, failing to recall, and taking out-of-date and inappropriately stored drugs were common home use medication errors. In our study, we tested errors related to recognition and perception of packages. Our findings suggest that perceptual recognition error can also be a factor in erroneous comparison of medication in real life.

Several studies have pointed to age as a factor in adverse medication usage, supporting our finding of significant age-related effects both on reaction time and errors. As Mira, Guilabert, et al. (2015) point out, the elderly are a vulnerable group due to higher risk of perceptual errors, a greater use of medication, and a greater risk of adverse effects of medication errors. It is therefore crucial to ensure that the probability of error is as low as possible.

In our study, highlighting substance name by providing a clear contrast band background for the name and dose and positioning this information in a dedicated place on the packages help users identify substances in generic packages faster and more accurately. The results of Experiment 2 suggest that the key to error reduction in the redesigned packages is the highlighted substance name placed in a high-contrast area (band or box). Randomizing the placement of this information on packages had less impact on reaction times or errors.

Our package redesign was based on general human factors guidelines. This study is, as far as we know, the first experimental test of those principles. One study by Gerhart et al. (2015) made a similar test program for package labels. The labels were redesigned in several versions following the Food and Drug Administration industry guidance for labeling and packaging (Aspden, Wolcott, Bootman, & Cronewlett, 2006). In that study, the alternate label designs included a three-dimensional tablet image; in addition, a color band highlighting the dosage strength increased contrast between the text, color bands, and the paper substrate. In a study with seven test users, they found that this alternate label led to fewer medication errors. In contrast to our study, they did not find any differences between the different designs related to reaction time or errors in reading brand name and dosage. Gerhart et al. performed realistic

user testing in a pharmaceutical environment but did not manipulate substance name or packaging in their study; only the medication labels were changed. Brand and dosage were their focal parameters; they studied the effect of a redesign only on pharmacists in a prescription-writing task, not on patients. Even so, their results match well with ours, indicating that highlighting of substance name and dose might be a key design factor in reducing medication errors.

Our study has some limitations. It was performed under controlled experimental conditions, and the MRT, although relevant for some aspects of medication recognition, is not a direct simulation of every aspect of users' handling of medication. Environmental factors, such as lighting conditions, and package design principles, such as tactile information and manipulation of package size, were not addressed in our experiments.

Despite these limitations, several aspects of our results are notable. First, because we used packages as they actually occur on the market as a control, the relative difference in recognition error suggests that our results might be relevant for more ecologically relevant contexts. Second, we manipulated two key variables that challenge the recommended design principles: package color and placement of substance name.

Color is a controversial topic in medication packaging design, and no unified recommendations exist on how to avoid confusion and reduce errors (Brandt, 2015). We did not set out to test explicitly the effect of color coding. Instead, we did test two possible causes of confusion. The first is when two packages contain the same substance but have different color schemes, which leads the users to erroneously conclude that the packages contain two different substances. We observed a dramatic improvement in this type of error for both user groups (from 69% to 19% for the elderly and 44% to 9% for the young) with the redesigned packages compared with original generic medications. There might be a perceptual component to these types of errors in addition to the known effects of motivation, lack of understanding of consequences, and misunderstandings of prescriptions. Thus it is the confusion between two identical medications that provides highest error rates with the original design.

As noted by Wolf et al. (2012), this is the type of error that is most prominent in overdose or "double dipping" and is therefore an extremely important factor to counter.

The second confusion type is when two packages contain different substances but the packages have similar color schemes, leadings users to erroneously conclude that the packages contain the same substance. In this case, the error rates with both original and redesigned packaging were smaller, and there was no difference between the original and redesigned packages. However, the elderly users produced more errors than the young users (elderly, 24%, and young, 5%). Because there was no significant difference between the two error types for the redesigned packages within user groups, these error rates might be a task-specific floor effect related to mental rotation. However, this finding should not be interpreted as a measure of expected medication errors in a real-life setting.

In our second experiment, we found that highlighting the substance name and dosage information improved reaction time and reduced errors. In contrast, displaying key information in a predictable place on the packages did not seem to be as important as highlighting it. This finding is in accordance with the Gerhart et al. (2015) study, wherein dosage strength was highlighted in a similar way as in our experiment. It is plausible that giving a prominent place to substance name and dose with a band of high-contrast color might be the key to better recognition of active substance in medications. We did not include elderly users in this experiment and therefore cannot rule out the possibility that the substance name placement might be more important for this group. Given the general effect of age, further studies of medication packaging design should include elderly users.

Age is a major concern for usage of OTC medications. As Mira, Guilabert, et al. (2015) point out, older patients' age means poorer cognitive states, greater number of medicines used, and increase in errors made. Our findings point in the same direction. Even corrected for age-related task effects, the elderly users were significantly more prone to error. The package redesign reduced performance differences between elderly and young users significantly. Because elderly use more strategies to perform the MRT (Dror et al., 2005), highlighting key information may better support more global strategies associated with age. To further investigate these assumptions, controlled field trials need to be preformed. The combination of field and laboratory experiments allows translation of sound human factors principles to evidence-based guidance (Estock et al., 2015).

CONCLUSION

A redesign of generic medication packages decreased recognition errors. The most prominent improvement occurred when users might be predisposed to believe that the packages contained two different substances when in fact they contained the same substance. The key factor in the redesigned packages was the highlighting of active substance name and dose on medication packages on a high-contrast background. These findings are consistent with design principles that guide perceptual recognition. The present study provides important insights into the understanding of the perceptual aspects of medication package design and suggests that minor changes in packaging design significantly improves users' ability to determine whether or not two different drugs contain the same active ingredient. Ecological field trials can provide insight into the importance of the suggested design changes relative to other causes of medication error.

KEY POINTS

- Elderly users make more mistakes recognizing generic medication than young users.
- The most prominent error was to wrongly assume that the matched and sample packages contained different substances (elderly, 55%, and young, 35%) when in fact they contained the same substance.
- Color confusion increased this type of error to 68% (elderly) and 43% (young) when package coloring indicated different medications.
- Redesign of packages with highlighted substance name on a high-contrast box or band significantly reduced this type of error.

REFERENCES

Aitsiselmi, Y., & Holliman, N. S. (2009). Using mental rotation to evaluate the benefits of stereoscopic displays. *SPIE Proceedings*, *7237*.

Arnheim, R. (2004). *Art and visual perception: A psychology of the creative eye* (2nd ed.). Berkeley: University of California Press.

Aspden, P., Wolcott, J., Bootman, L., & Cronenwett, L. R. (2006). *Preventing medication errors*. Washington, DC: National Academic Press.

Bailey, S. C., Pandit, A. U., Curtis, L., & Wolf, M. S. (2009). Availability of Spanish prescription labels: A multi-state pharmacy survey. *Medical Care*, *47*, 707–710.

Baltes, P. B., & Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult lifespan: A new window to the study of cognitive aging? *Psychology and Aging*, *12*, 12–21.

Barfield, W., Sandford, J., & Foley, J. (1988). The mental rotation and perceived realism of computer-generated three-dimensional images. *International Journal of Man–Machine Studies*, *29*, 669–684.

Berman, A. (2004). Reducing medication errors through naming, labeling, and packaging. *Journal of Medical Systems*, *28*, 9–29.

Brandt, J. D. (2015). Human factors and ophthalmic drug packaging: Time for a global standard. *Ophthalmology*, *122*, 2368–2370.

Chafac, M. M. N., & Chan, A. H. S. (2012). Age and comprehension of written medical information on drug labels. In K. M. Stanney & K. S. Hale (Eds.), *Advances in cognitive engineering and neuroergonomics* (pp. 207–217). Boca Raton, FL: CRC Press.

Cohen, M. R. (2000). *Medication errors*. London, UK: Jones and Bartlett.

Cooper, L. A., & Podgorny, P. (1976). Mental transformations and visual comparison processes: Effects of complexity and similarity. *Journal of Experimental Psychology: Human Perception and Performance*, *2*, 503–514.

Dror, I. E., Schmitz-Williams, I. C., & Smith, W. (2007). Older adults use mental representations that reduce cognitive load: Mental rotation utilizes holistic representations and processing. *Experimental Aging Research*, *31*, 409–420.

Estock, J. L., Murray, A. W., Mizah, M. T., Mangione, M. P., Goode, J. S., & Eibling, D. E. (2015). Label design affects medication safety in an operating room crisis: A controlled simulation study. *Journal of Patient Safety*. Advance online publication.

European Parliament & Council of the European Union. (2001). *Directive 2001/83/EC of the European Parliament and of the Council*. Brussels, Belgium: Author.

Filik, R., Purdy, K., Gale, A., & Gerrett, D. (2006). Labeling of medicines and patient safety: Evaluating methods of reducing drug name confusion. *Human Factors*, *48*, 39–47.

Fraeyman, J., Peeters, L., Van Hal, G., Beutels, P., De Meyer, G. R., & De Loof, H. (2015). Consumer choice between common generic and brand medicines in a country with a small generic market. *Journal of Managed Care & Specialty Pharmacy*, *21*, 288–296.

Gerhart, J. M., Spriggs, H., Hampton, T. W., Hoy, R. M. B., Strochli, A. Y., Proulx, S., & Goethius, D. B. (2015). Applying human factors to develop an improved package design for (Rx) medication drug labels in a pharmacy setting. *Journal of Safety Research*, *55*, 177–184.

Glisky, E. L. (2007). Changes in cognitive function in human aging. In D. R. Riddle (Ed.), *Brain aging: Models, methods, and mechanisms* (pp. 3–20). Boca Raton, FL: CRC Press/Taylor & Francis.

Håkonsen, H., Eilertsen, M., Borge, H., & Toverud, E. L. (2009). Generic substitution: Additional challenge for adherence in hypertensive patients. *Current Medical Research and Opinion*, *10*, 2515–2521.

Håkonsen, H., & Toverud, E. L. (2010). Special challenges for drug adherence following generic substitution in Pakistani immigrants living in Norway. *European Journal of Clinical Pharmacology*, *67*, 193–201. doi:10.1007/s00228-010-0960-9

Hancock, M., Carpendale, S., & Cockburn, A. (1988). Shallow-depth 3D interaction: Design and evaluation of one-, two- and three-touch techniques. *International Journal of Man–Machine Studies*, *29*, 669–684.

Helen Hamlyn Research Center & National Safety Agency. (2007). *Design for patient safety: A guide to the graphic design of medication packaging* (2nd ed.). Retrieved from http://www.nrls.npsa.nhs.uk/resources/collections/design-for-patient-safety/?entryid45=63053

Holden, R. J., Schubert, C. C., & Mickelson, R. S. (2015). The patient work system: An analysis of self-care performance barriers among elderly heart failure patients and their informal caregivers. *Applied Ergonomics*, *47*, 133–150.

Kenagy, J. W., & Stein, G. C. (2001). Naming, labeling, and packaging of pharmaceuticals. *American Journal of Health-System Pharmacy*, *58*, 2033–2041.

Kjoenniksen, I., Lindbaek, M., & Granas, A. G. (2006). Patients' attitudes towards and experiences of generic drug substitution in Norway. *Pharmacy World and Science*, *28*, 284–289.

Kongkaew, C., Noyce, P. R., & Ashcroft, D. M. (2008). Hospital admissions associated with adverse drug reactions: A systematic review of prospective observational studies. *Annals of Pharmacotherapy*, *42*, 1017–1025.

Mira, J. J., Guilabert, M., César, I., Fernández, C., Vicente, A., Orozco-Beltrán, D., & Gil-Guillen, V. F. (2015). Use of QR and EAN-13 codes by older patients taking multiple medications for a safer use of medication. *International Journal of Medical Informatics*, *84*, 406–412.

Mira, J. J., Lorenzo, S., Guilabert, M., Navarro, I., & Pérez-Jover, V. (2015). A systematic review of patient medication error on self-administering medication at home. *Expert Opinion on Drug Safety*, *14*, 815–838.

Peters, M., & Battista, C. (2007). Applications of mental rotation figures of the Shepard and Metzler type and description of a mental rotation stimulus library. *Brain and Cognition*, *66*, 260–264.

Pinker, S., & Finke, R. (1980). Emergent two-dimensional patterns in depth. *Journal of Experimental Psychology: Human Perception and Performance*, *6*, 244–264.

Rajashekar, U., Cormack, L. K., & Bovik, A. C. (2002). Visual search: Structure from noise. In *Proceedings of the Eye Tracking Research & Applications Symposium* (pp. 119–123). New York, NY: ACM.

Shepard, R. N., & Metzler, J. (1971). Mental rotation: Effects of dimensionality of objects and type of tasks. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 3–11.

Shepard, R. N., & Cooper, L. A. (1982). *Mental images and their transformations*. Cambridge, MA: MIT Press/Bradford Books.

Ward, J., Buckle, P., & Clarkson, J. P. (2010). Designing packaging to support the safe use of medicines at home. *Applied Ergonomics*, *41*, 682–694.

Weiss, S. (2009). Compliance packaging for over-the-counter drug products. *Journal of Public Health*, *17*, 155–164.

Wolf, M. S., Davis, T. C, Shrank, W., Rapp, D. N., Connor, U. M., & Parker, R. M. (2007). To err is human: Patient misinterpretations of prescription drug label instructions. *Patient Education and Counseling*, *67*, 293–300.

Wolf, M. S., King, J., Jacobson, K., DiFransesco, L., Bailey, S. C., Mullen, R., … Parker, R. M. (2012). Risk of unintentional overdose with non-prescription acetaminophen products. *Journal of General Internal Medicine*, *27*, 1587–1593.

Yin, H. S., Mendelsohn, A. L., Wolf, M. S., Parker, R. M., Fierman, A., van Schaick, L., Bazan, I. S., Kline, M. D., & Dreyer, B. P. (2010). Parents' medication administration errors: Role of dosing instruments and health literacy. *Archives of Pediatric and Adolescence Medicine*, *164*, 181–186.

Tor Endestad is an associate professor in cognitive neuroscience at the Department of Psychology, University of Oslo, where he received a PhD in psychology in 2002. He is a former human factors associate at the Organisation for Economic Co-operation and Development Halden Reactor Project and now head of research at the Department of Psychology. He coheads the FRONT Neurolab at the University of Oslo.

Laura A. Wortinger is a PhD research fellow in neuroscience at the Department of Psychology, University of Oslo. She has a master's degree in cognitive neuroscience from University of Oslo. Her PhD is on functional brain correlates of fatigue.

Steinar Madsen is a medical director at the Norwegian Medicines Agency. He graduated with a MD from the University of Oslo in 1981. He is a specialist in internal medicine and cardiology and works part-time as a consultant in cardiology. Since 2001, he is a member and previously chairman of the committee for generic substitution at the agency.

Sigurd Hortemo is a senior medical advisor at the Norwegian Medicines Agency. He graduated with an MD from the University of Oslo in 1980. He is a specialist in family medicine and community medicine and works part-time as a family doctor in Oslo.