Labeling of Medicines and Patient Safety: Evaluating Methods of Reducing Drug Name Confusion

Ruth Filik, University of Glasgow, Glasgow, United Kingdom, and **Kevin Purdy, Alastair Gale,** and **David Gerrett,** University of Derby, Derby, United Kingdom

Objective: We report three experiments evaluating the proposal that highlighting sections of drug names using uppercase ("tall man") lettering and/or color may reduce the confusability of similar drug names. **Background:** Medication errors commonly involve drug names that look or sound alike. One potential method of reducing these errors is to highlight sections of names on labels in order to emphasize the differences between similar products. **Method:** In Experiments 1 and 2, participants were timed as they decided whether similar name pairs were the same name or two different names. Experiment 3 was a recognition memory task. **Results:** Results from Experiments 1 and 2 showed that highlighting sections of words using tall man lettering can make similar names easier to distinguish if participants are aware that this is the purpose of the intervention. Results from Experiment 3 suggested that tall man lettering and/or color does not make names less confusable in memory but that tall man letters may increase attention. **Conclusion:** These findings offer some support for the use of tall man letters in order to reduce errors caused by confusion between drug products with look-alike names. **Application:** The use of tall man letters could be applied in a variety of visual presentations of drug names – for example, by manufacturers on packaging, labeling, and computer software, and in pharmacies on shelf labels. Additionally, this paper demonstrates two meaningful behavioral measures that can be used during product design to objectively assess confusability of packaging and labeling.

INTRODUCTION

Recent reports cite medical error as being one of the leading causes of death in developed countries, with errors involving medication being among the most common of these mistakes (e.g., Kohn, Corrigan, & Donaldson, 2000). A medication error can be defined as "any event that could cause or lead to a patient's receiving inappropriate drug therapy or failing to receive appropriate drug therapy. The event could occur at any point from the decision to initiate therapy to the point at which the patient received the medication" (Edgar, Lee, & Cousins, 1994, p. 1336). Similar-looking commercial labeling and packaging are common causes of medication error, often from the use of nearly identical packaging for two separate items. Edgar et al. (1994) noted that the incidents most commonly reported to the U.S. Pharmacopeia Medication Errors Reporting Program between August 1991 and April 1993 involved problems with the product (e.g., similar packaging or incomplete labeling).

Errors can also result from products having similar names. One recent observational study suggested that wrong drug errors per prescription occur at a rate of 0.13% (Flynn, Barker, & Carnaham, 2003). Voluntary report figures put the proportion of errors that cite name confusion as the cause as high as 25% (see, e.g., Hoffman & Proulx, 2003; Kenagy & Stein, 2001). Mistakes can occur because of similarities between names, whether proprietary (i.e., brand, trademark) or nonproprietary (i.e., generic). Products may be given similar brand names so that the value invested in one product is easily transferred to another. Generic names must be meaningful to health care professionals and patients, resulting in drugs that share an indication, mechanism of action, or chemical constituent often intentionally being given the

Address correspondence to Ruth Filik, Department of Psychology, University of Glasgow, 58 Hillhead St., Glasgow G12 8QB, UK; r.filik@psy.gla.ac.uk. ***HUMAN FACTORS,*** Vol. 48, No. 1, Spring 2006, pp. 39–47. Copyright © 2006, Human Factors and Ergonomics Society. All rights reserved.

Downloaded from hfs.sagepub.com at Swansea University on February 16, 2015

same prefix or suffix. For example, 35% of generic drug names end in *–ne* (Lambert, Chang, & Lin, 2001a).

Strategies for reducing errors involving drug name confusion must consider both preventing the approval of new names that may be confused with existing names and dealing with existing confusable name pairs. Lambert and colleagues (e.g., Lambert, Chang, & Lin, 2001b; Lambert, Donderi, & Senders, 2002; Lambert, Lin, Chang, & Gandhi, 1999) argued that an effective method of preventing potentially confusable names from being approved would involve utilizing computer software to screen all proposed names for similarity with existing names (using objective measures of similarity as outlined in Lambert et al., 1999). These objective measures could be combined with subjective opinions from panels of experts.

To prevent confusion errors between existing names, the use of bar codes, automated dispensing, and nonalphabetical storage of drug products has been suggested (M. R. Cohen, 1999). Another possible solution is to change the appearance of names on product labels, computer screens, shelf labels, and so forth. For instance, M. R. Cohen (1999) suggested that it is much easier to differentiate between "DOBUTamine" and "DOPamine" than between "dobutamine" and "dopamine." The Food and Drug Administration (FDA) Name Differentiation Project implemented this idea in 2001. Following FDA recommendations, the Office of Generic Drugs requested that manufacturers of 16 look-alike name pairs should voluntarily revise the appearance of their established names in order to minimize medication errors. Manufacturers were encouraged to produce labeling that visually differentiates the established names with the use of "tall man" (uppercase) letters (see http://www.fda.gov/cder/drug/MedErrors/nameDiff.htm for a list of the names involved and the recommended revisions). Following these recommendations, a manufacturer would capitalize the letters in the drug name that are different from the name's stem. For example, "chlorpropamide" and "chlorpromazine" both have "chlorpro" at the beginning, so the remaining letters of each name would be capitalized and perhaps printed in a different color – for example, "chlorproMAZINE" and "chlorproPAMIDE."

Typeface characteristics are an important factor in label design. Much recent research has concentrated on the effects of print characteristics on the readability of warnings and other consumer information (see Rogers, Lawson, & Rousseau, 2000, for a review of warning research). For example, studies have shown that legibility is reduced when characters are horizontally compressed (Watanabe, 1994), when thinner type widths are used (Young, Laughery, & Bell, 1992), and when spacing between letters is reduced (Anderton & Cole, 1982). Other studies have demonstrated that increasing the vertical spacing between text aids comprehension (Hartley, 1999), increasing font size from 8 to 10 points and using bold font increases perceived readability (Silver & Braun, 1993), and conspicuously printed warnings with larger, wider-stroke print and orange highlighting are recalled better than less conspicuous warnings with smaller, thinner, nonhighlighted print (Young & Wogalter, 1990).

Some recent research has focused on the labeling of medication. Smither and Braun (1994) found that younger and older adults took longer to read labels printed in 9-point font than in 12- or 14-point font. Labels printed in 14-point font were judged easier to read than those in 9- and 12-point fonts. Wogalter and Vigilante (2003) found that older consumers may be unable to acquire information in the "fine" (4-point) print frequently found in various kinds of product literature.

Despite this interest in readability, relatively little research has studied confusion between products with similar names and/or labeling. It is widely accepted that similarity has adverse effects on a number of cognitive and perceptual processes (see, e.g., Reason, 1990, for a discussion of the cognitive processes underlying human error). The purpose of this paper is to evaluate possible methods of reducing confusion between similarly named drug products, specifically, the use of tall man and/or colored lettering to emphasize the differences between products with similar names.

EXPERIMENT 1

The aim of Experiment 1 was to investigate the effect of tall man letters on perceptual similarity. Participants were given a "same/different" judgment task, in which they were presented with a pair of drug names on a computer screen and had to indicate, as quickly and as accurately as possible, whether the two names were the same name presented twice or were two different names. The

Downloaded from hfs.sagepub.com at Swansea University on February 16, 2015

task was chosen to represent a situation in which people are faced with similarly named products that are placed next to each other on a shelf. It was hypothesized that if highlighting the different parts of similar-looking names on drug packaging (e.g., "chlorproMAZINE" and "chlorproPAMIDE") makes them easier to distinguish, then participants should be faster to indicate that a pair of similar names are different from each other when the names contain tall man letters than when they are presented all in lowercase.

Method

Participants. Twenty staff and students from the University of Derby (who were not health care professionals) participated.

Materials and design. Eighty pairs of generic drug names were selected as materials, 16 of which had been recommended for the FDA Name Differentiation Project. The remaining 64 were selected from Davis, Cohen, and Teplitsky (1992), who published a list of 645 similar-looking drug names. These 64 pairs were selected on the grounds that the names in each pair did not vary in length from each other by more than one character, so participants could not use word length as a cue. Similarity was also assessed using an objective measure. Orthographic similarity was measured using a *bigram* method, with one space added to the beginning and end of each word (as used in Lambert et al., 2001b; see also Lambert et al., 1999, for a discussion of this and other measures of similarity). To compute the bigram similarity score, it was first necessary to split the words into two-letter sequences. For example, the bigrams for "dobutamine" were {_d, do, ob, bu, ut, ta, am, mi, in, ne, e_} and those for "dopamine" were {_d, do, op, pa, am, mi, in, ne, e_}. A similarity score between 0 and 1 was then computed using the Dice coefficient, orthographic similarity = $2C/(B + A)$, in which A is the number of bigrams in the first word, B is the number of bigrams in the second word, and C is the number of bigrams that occur in both words. "Dobutamine" and "dopamine" share seven bigrams {_d, do, am, mi, in, ne, e_}, giving a bigram similarity score of $(2 \times 7)/(11 + 9) = .7$. The mean bigram similarity score for the remaining 64 pairs used as materials was .6 ($SD = .1$).

The program to present the stimuli was written in Psychology Software Tools E-Prime Version 1.1. The names were presented on a 17-inch (43-cm) CRT monitor in Arial font as black text on a white background (as medication labels commonly comprise black sans serif font on a white background) at a screen resolution of 640 × 480. The shortest name used subtended 3.3° and the longest name subtended 10.3° of visual angle horizontally to the participant. The height of the uppercase letters was 0.8°, and the names were separated by a gap of 2.1°.

Name pairs were presented with either both names in lowercase letters or both names having sections in tall man letters. They could consist of two different names or of the same name presented twice. Four stimulus files were created so that each name pair could be presented in each of the following four conditions without participants seeing any individual name pair more than once (resulting in 20 pairs in each condition in each file).

lowercase + same
chlorpromazine chlorpromazine
lowercase + different
chlorpromazine chlorpropamide
tall man + same
chlorproMAZINE chlorproMAZINE
tall man + different
chlorproMAZINE chlorproPAMIDE

There were two within-participants independent variables: similarity (names were either the same or different) and letter style (lowercase or tall man). The dependent variable was the time taken by the participant to indicate whether the names were the same or different. In cognitive psychology, it is conventional to expect a tight link between response time measures and task difficulty; therefore any influence of tall man lettering on discriminability should be reflected in response times. Error rate was not used, because participants had as long as they needed to make their decision.

Procedure. Participants read a set of instructions on the computer screen. Their task was to decide, as quickly and as accurately as possible, whether the pairs of names that were presented to them comprised the same name presented twice or two different names. They indicated their response via two keys (marked "same" and "different") on a serial response box. Participants fixated a cross at the center of the screen that disappeared after 1 s and was replaced by the name pair, which remained on the screen until the participant pressed

Downloaded from hfs.sagepub.com at Swansea University on February 16, 2015

one of the keys. Four practice trials preceded the 80 experimental trials. The experiment took approximately 5 min in total.

Results

Response time was taken as a measure of how easy the name pairs were to distinguish. Mean response times are shown in the top portion of Table 1.

Data were analyzed using two 2 (similarity) × 2 (letter style) repeated measures analyses of variance (ANOVAs). One ANOVA treated participants (F_1) as the random variable, and the other treated items (F_2) as the random variable (Clark, 1973). Partial eta-squared (η_p^2) was calculated as a measure of effect size, with values of .01, .06, and .14 characterizing small, medium, and large effect sizes, respectively (J. Cohen, 1988). There was a significant effect of similarity, $F_1(1, 19) = 22.83$, $p < .0005$, $\eta_p^2 = .55$; $F_2(1, 79) = 123.95$, $p < .0005$, $\eta_p^2 = .61$., with longer response times when the two names were the same than when they were different. There was no effect of letter style, $F_1(1, 19) = .12$, $p > .05$, $\eta_p^2 = .01$; $F_2(1, 79) = .08$, $p > .05$, $\eta_p^2 < .01$, and no interaction, $F_1(1, 19) = .62$, $p > .05$, $\eta_p^2 = .03$; $F_2(1, 79) = .20$, $p > .05$, $\eta_p^2 < .01$, demonstrating that participants could discriminate between the tall man names no faster than they could discriminate between the lowercase names.

Discussion

Response times were longer when name pairs were the same than when they were different. This is probably because participants could judge name pairs as being different on detection of a single difference between the names. In order to judge them as being the same, however, they had to scan the entire name. Importantly, results from Experiment 1 indicate that name pairs containing tall man letters are no easier to distinguish from each other than are names that are presented in lowercase. However, it is likely that in a real-life situation, people would be aware that the uppercase sections of the names were particularly important and would therefore adopt a different strategy based on this knowledge. This possibility was investigated in Experiment 2.

TABLE 1: Mean Response Times (*SD*) by Condition for Experiments 1 and 2

	Response Time (ms)	
	Lowercase	Tall Man
Experiment 1		
Same	2214 (979)	2223 (894)
Different	1752 (650)	1713 (578)
Experiment 2		
Same	1571 (594)	1451 (448)
Different	1521 (540)	1320 (364)

EXPERIMENT 2

Method

Twenty staff and students from the University of Derby (non–health care professionals, who had not taken part in Experiment 1) participated. The materials and design were identical to those of Experiment 1. The procedure was also identical, except that participants were given the additional instruction that tall man letters were used in an attempt to make similar names less confusable with each other.

Results

Once again, response time was taken as a measure of how easy the name pairs were to distinguish. Mean response times are shown in the bottom portion of Table 1.

Data were analyzed using two 2 (similarity) × 2 (letter style) repeated measures ANOVAs. One ANOVA treated participants as the random variable (F_1), and the other treated items as the random variable (F_2). Partial eta-squared (η_p^2) was calculated as a measure of effect size. There was a significant effect of similarity, $F_1(1, 19) = 13.14$, $p < .005$, $\eta_p^2 = .41$; $F_2(1, 79) = 27.40$, $p < .005$, $\eta_p^2 = .26$, with longer response times for names that were the same than for names that were different. There was also a significant effect of letter style, $F_1(1, 19) = 4.71$, $p < .05$, $\eta_p^2 = .20$; $F_2(1, 79) = 7.23$, $p < .01$, $\eta_p^2 = .08$, showing shorter response times for the tall man name pairs than for the lowercase names, with η_p^2 indicating a medium to large effect size. There was no interaction, $F_1(1, 19) = 1.30$, $p > .05$, $\eta_p^2 = .06$; $F_2(1, 79) = 1.63$, $p > .05$, $\eta_p^2 = .02$.

Discussion

Results from Experiment 2 suggest that if participants are aware of the purpose of tall man

Downloaded from hfs.sagepub.com at Swansea University on February 16, 2015

lettering, then they can discriminate more easily between name pairs with tall man lettering than between lowercase name pairs. Taking results from Experiment 1 into account, this would suggest that tall man lettering does not intrinsically make look-alike name pairs easier to differentiate but can be effective if participants are aware of its purpose. Tall man letters are not the only way in which sections of similar names can be highlighted. There are numerous other possibilities, including the use of color, boldface, or italics. Some of these alternatives are considered in Experiment 3.

EXPERIMENT 3

To narrow down the number of possible manipulations, 20 University of Derby staff and students were asked for their subjective opinions on a number of different ways to highlight text. Participants were given a piece of paper with the drug names "vincristine" and "vinblastine" printed on it seven times, with each pair of names in a different format. The names were printed normally ("vincristine," "vinblastine"), with the distinguishing section in tall man letters ("vinCRIStine," "vinBLAStine"), in color ("vin[cris]tine," "vin[blas]tine," in which square brackets indicate the letters printed in red), in italics ("vin*cris*tine," "vin*blas*tine"), in bold ("vin**cris**tine," "vin**blas**tine"), underlined ("vin<u>cris</u>tine," "vin<u>blas</u>tine"), and in larger lowercase font ("vincristine," "vinblastine"). Participants were asked to rate each name pair on how confusable they looked on a scale from 1 (*not at all confusable*) to 7 (*extremely confusable*). Mean confusability ratings are shown in Table 2.

A repeated measures analysis of variance showed a significant difference between the conditions, $F(1, 19) = 13.57, p < .001$. Pairwise comparisons revealed that names in the control condition were judged to be significantly more confusable than names in all of the other conditions ($ps < .05$). Names printed with sections in color were judged to be significantly less confusable than all of the other manipulations (all $ps < .05$). This is perhaps not surprising, as color is known to be a highly salient visual surface feature and is an effective means of engaging visual attention (e.g., Davidoff, 1991). Therefore in Experiment 3 we investigated whether using color, or a combination of color and tall man lettering, would be any more effective than using tall man letters alone.

We decided to use a difficult task, a recognition memory task, so that error rates could be measured rather than response times, as this may more accurately reflect the processes in which we are interested (i.e., the likelihood of errors being made). Recognition memory tasks comprise a study phase and a test phase. During the study phase, participants are presented with the material to be remembered (in this case, a list of drug names). These names are then presented to the participant again during the test phase, when they are mixed in with distractor words. The participant's task is to indicate which words were present in the study list and which were not (i.e., are distractors). It is well established that similarity can affect recognition memory. Distractor names that are similar to study words orthographically (have a similar arrangements of letters), phonologically (are similar sounding), or semantically (have similar meanings) are falsely recognized more frequently than are unrelated distractor items (e.g., Anderson, Bothell, Lebiere, & Mantessa, 1998; Dale & Baddeley, 1962).

This task is analogous to the real-world situation in which a patient or health care professional has committed the name of a drug to memory and, subsequently, has to search for that drug among other products on a shelf, some of which may have names similar to that of the drug he or she is looking for. A false recognition error in this setting could lead to the incorrect selection of a drug that has a name similar to that of the intended drug (as noted in Lambert et al., 2001b).

Recently, research has been carried out to determine the effects of using orthographically and phonologically similar distractors on the recogni-

TABLE 2: Confusability Ratings by Condition for Experiment 3

	Color	Tall Man	Larger	Bold	Underline	Italics	Control
Mean rating	1.90	2.55	2.60	3.40	3.45	4.40	5.40
SD	(0.98)	(0.94)	(1.27)	(1.14)	(1.50)	(1.05)	(1.35)

Downloaded from hfs.sagepub.com at Swansea University on February 16, 2015

tion of drug names. Lambert et al. (2001b) expected false positive recognition errors to increase as similarity between target and distractor names increased. They used pairs of names at increasing levels of similarity as determined by a bigram similarity test (described earlier). Logistic regression was carried out on the number of false positive recognition errors. For both pharmacists and university students, increased orthographic and phonological similarity was associated with a significant increase in the log odds of a false recognition. The results suggested that excessive similarity between drug names posed a risk of increased errors.

If tall man letters and/or color make similar drug names less confusable with each other, we would predict that in a recognition memory experiment, fewer distractors would be incorrectly identified as targets (i.e., there would be fewer false positive errors) for names that contain tall man and/or colored letters than for names that do not.

Method

Participants. Twenty-eight staff and students from the University of Derby (non–health care professionals, who had not taken part in any of the previous experiments) participated.

Materials and design. Sixty pairs of generic drug names were selected as materials. One name from each pair was used in the study phase of the experiment, and the other names were used as distractor items. Sixteen of these were pairs of drug names that had been recommended for the FDA Name Differentiation Project. The remaining 44 pairs were selected from Davis et al. (1992). As with Experiments 1 and 2, an objective test of similarity (the bigram method with one space added to the beginning and end of each word) was also adopted when selecting these 44 materials. Only name pairs with a similarity score of .5 or above were selected (mean = .7, *SD* = .1). Highly similar name pairs were selected as materials as it is these names that would benefit from the proposed interventions.

The program to present the stimuli was written in Psychology Software Tools E-Prime Version 1.1. The words were presented on a 17-inch (43-cm) CRT monitor in Arial font as black or black-and-red text on a white background at a screen resolution of 640 × 480. The stimuli were created in Corel Photo-Paint. The shortest name subtended 4.2° of visual angle and the longest name subtended 9.8° of visual angle horizontally to the participant. The height of the uppercase letters was 0.8°.

Names were presented either in lowercase or with sections in tall man letters, and they could consist of black text alone or of black-and-red text. (In the following examples, square brackets are used to indicate the letters presented in red.) Four stimulus files were created so that each name pair could be presented in each of the following four conditions without participants seeing each name more than once (resulting in 30 names in each condition in the test phase):

lowercase + not colored:
chlorpromazine chlorpropamide
lowercase + colored:
chlorpro[mazine] chlorpro[pamide]
tall man + not colored:
chlorproMAZINE chlorproPAMIDE
tall man + colored:
chlorpro[MAZINE] chlorpro[PAMIDE]

The four conditions were rotated in a Latin square across each experimental file, but presentation was randomized within each study-test block. There were two independent variables: letter style (lowercase or tall man) and color (colored or not colored). The dependent variables were the overall number of errors, the number of false negative errors (i.e., when a name was reported as being a distractor when it had in fact been present in the study list), and the number of false positive errors (i.e., when a name was reported as being from the study list when it was in fact a distractor) in each condition. False positive errors correspond to medication errors.

Procedure. Participants read a set of printed instructions. The task consisted of 1 practice block and 12 experimental blocks. Each block comprised a study phase and a test phase. During the study phase, participants fixated a cross at the center of the screen. This cross was then replaced by 5 names that appeared on the screen 1 at a time. Each name remained on the screen for 2 s. After the 5th name had been shown, participants were presented with a message indicating that the test phase was about to begin, along with a reminder of the task. During the test phase, 10 names were presented, 1 at a time. These names were the 5 names that had appeared during the study phase

Downloaded from hfs.sagepub.com at Swansea University on February 16, 2015

(each of which consisted of 1 of the names from a confusable name pair) and 5 distractor items (the other name from the same 5 name pairs). The participants' task was to indicate, for each name on the test list, whether or not it had appeared in the study list. They had to make this decision as quickly and as accurately as possible and indicate their response via two keys on a response box (one key marked "old," to indicate that the name had been present in the study phase, and one marked "new," to indicate that it had not). The experiment took approximately 10 min in total.

Results

Three measures of recognition accuracy are reported: the overall number of errors, the number of false negative errors, and the number of false positive errors made in each condition. Means for each kind of error are shown in Table 3.

Two 2 (letter style) × 2 (color) repeated measures ANOVAs (F_1 and F_2) were performed on the overall number of errors, the number of false negative errors, and the number of false positive errors. Partial eta-squared (η_p^2) was calculated as a measure of effect size. For the overall number of errors, there was a significant effect of letter style, $F_1(1, 27) = 8.16$, $p < .01$, $\eta_p^2 = .23$; $F_2(1, 59) = 5.29$, $p < .05$, $\eta_p^2 = .08$, showing fewer overall errors for names containing tall man letters than for names that were all in lowercase, with partial eta-squared indicating a medium to large effect size. Color did not have a significant effect on the overall number of errors, $F_1(1, 27) = .26$, $p > .05$, $\eta_p^2 < .01$; $F_2(1, 59) = .95$, $p > .05$, $\eta_p^2 = .02$, and there was no interaction between color and letter style, $F_1(1, 27) = .14$, $p > .05$, $\eta_p^2 < .01$; $F_2(1, 59) = .05$, $p > .05$, $\eta_p^2 < .01$. For false negative errors, there was a significant effect of letter style, $F_1(1, 27) = 6.63$, $p < .05$, $\eta_p^2 = .20$; $F_2(1, 59) = 4.49$, $p < .05$, $\eta_p^2 = .07$, showing fewer false negative errors for names containing tall man letters than for names that were all in lowercase, with a medium to large effect size. There was no effect of color, $F_1(1, 27) = .41$, $p > .05$, $\eta_p^2 = .02$; $F_2(1, 59) = .88$, $p > .05$, $\eta_p^2 = .02$, and no interaction, $F_1(1, 27) = .88$, $p > .05$, $\eta_p^2 = .03$; $F_2(1, 59) = 1.05$, $p > .05$, $\eta_p^2 = .02$. For false positive errors, there was no effect of letter style, $F_1(1, 27) = 1.15$, $p > .05$, $\eta_p^2 = .04$; $F_2(1, 59) = 1.85$, $p > .05$, $\eta_p^2 = .03$, or color, $F_1(1, 27) = .003$, $p > .05$, $\eta_p^2 < .01$; $F_2(1, 59) = .22$, $p > .05$, $\eta_p^2 < .01$, and no interaction, $F_1(1, 27) = .57$, $p > .05$, $\eta_p^2 = .02$; $F_2(1, 59) = 1.50$, $p > .05$, $\eta_p^2 = .02$.

Discussion

Experiment 3 showed that overall, drug names containing tall man letters were recognized more accurately than names printed all in lowercase. Printing sections of the names in color did not aid recognition, and there was no significant benefit of a tall man + color combination over using tall man letters alone. However, there was no significant difference in the number of false positive errors across conditions, suggesting that neither tall man lettering nor color made similar names less confusable with each other in memory (as they did not reduce the number of distractor names that were falsely recognized as targets). Whether or not manipulations of letter style would be expected to have an effect in this task would depend on how these words are mentally represented (e.g., see Johnston, McKague, & Pratt, 2004, for a recent discussion of this issue).

False negative error rates would suggest that tall man lettering simply made the names more memorable, probably attributable to tall man letters increasing attention to those names. It is perhaps surprising that color did not also have this effect, as research has shown that color is a highly salient visual attribute that increases attention, thereby improving recognition memory. For example, Wichmann, Sharpe, and Gegenfurtner (2002) found improved performance for recognition memory of naturally colored images over

TABLE 3: Mean (*SD*) Number of Errors by Condition in Experiment 3

	Lowercase		Tall Man	
Kind of Error	Black	Colored	Black	Colored
Overall	10.68 (3.76)	10.21 (3.28)	9.32 (3.76)	9.18 (3.09)
False negative	6.21 (2.74)	5.57 (2.30)	5.04 (2.74)	5.00 (2.19)
False positive	4.43 (2.46)	4.61 (2.20)	4.29 (2.64)	4.07 (1.88)

Downloaded from hfs.sagepub.com at Swansea University on February 16, 2015

black-and-white images, largely attributable to a decreased false negative error rate and not to differences in the false positive error rate, similar to our pattern of findings for tall man letters. This result would suggest that using tall man lettering is a more effective strategy than using color in increasing attention to these high-risk names. It should be noted that it would also be more appropriate for members of the population who are color blind.

It could be argued that the memory benefit for tall man lettering was attributable to participants adopting a strategy whereby they simply remembered the section of the name that was presented in uppercase letters, which would be easier than remembering the whole word. However, if this was the case, there was no reason they should not adopt a similar strategy for the names that had sections in color, but there was no observed memory benefit for the color condition.

GENERAL DISCUSSION

The aim of this paper was to evaluate the use of tall man lettering and/or color as an intervention to reduce medication errors involving drugs with confusable names. Results from Experiments 1 and 2 suggest that tall man lettering can make similar drug names easier to distinguish as long as people are aware of its purpose (as they would be in a real-life situation). Experiment 3 showed a benefit for tall man lettering in recognition memory for names. The nature of the benefit (i.e., fewer false negative errors rather than fewer false positive errors) would suggest that the tall man intervention increases attention to high-risk drug names rather than actually making them less confusable with each other in memory. The results would also suggest that tall man lettering is more effective than color at increasing attention in this situation and that combining tall man lettering with color gives no added benefit over tall man lettering alone. It is interesting to note that participants subjectively rated color as the most effective way of making names less confusable, which does not match the findings of the memory task. This is perhaps not surprising, however. Wogalter and Vigilante (2003) observed that research in psychology and human factors frequently shows a mismatch between measures of performance and subjective judgments, especially when the two measures may be tapping into different processes (e.g., Wogalter, Begley, Scancorelli, & Brelsford, 1997; Wogalter & Dingus, 1999).

Although our studies have provided interesting results, there are a number of limitations. The current research used participants who were not health care professionals. Therefore participants may not have been familiar with all of the names used in the study, and it is possible that these results may not generalize to people who are more familiar with drug names. However, non–health care professionals (e.g., patients choosing drug products in a commercial pharmacy) also make errors involving medication. Importantly, Lambert et al. (2001b) used a recognition memory task, similar to Experiment 3 reported here, to study the effect of orthographic similarity on the likelihood of participants making false positive errors; they found the same pattern of results for pharmacists as they did for university students.

It should be noted that similarity between drug products includes, but is not limited to, look-alike and sound-alike names. The problem of similarity can be compounded by other factors that may contribute to error. Examples include poor handwriting on prescriptions, similar packaging and labeling (particularly if the drugs are made by the same manufacturer), and similarly named drugs that also share strengths and indications. Other influences on the likelihood of error include environmental factors such as noise (Flynn et al., 1996), lighting conditions (Buchanan, Barker, Gibson, Jiang, & Pearson, 1991), and distractions (Flynn et al., 1999). The aim of the current studies was to examine name confusion error in a controlled laboratory environment, and therefore they necessarily excluded other influences, whereas it should be noted that errors in real life are generally caused by the interaction of several contributory factors. Despite these limitations, the current studies suggest that tall man letters can make similar names less confusable perceptually and can increase attention to high-risk drug names.

Implications for Patient Safety

These findings offer support for the use of tall man letters in order to reduce errors caused by confusion between drug products with look-alike names. This intervention could be applied in a variety of visual presentations of drug names – for example, by manufacturers on packaging, labeling, and computer software; and in pharmacies

Downloaded from hfs.sagepub.com at Swansea University on February 16, 2015

on shelf labels. Furthermore, it has been argued that it is important to incorporate human factors concepts such as differentiation and unambiguous communication into the design of medication-related objects and procedures in order to reduce error (Kenagy & Stein, 2001). This paper demonstrates two meaningful behavioral measures that can be used during product design to objectively assess confusability of packaging and labeling.

ACKNOWLEDGMENTS

The authors would like to thank Kevin Paterson for his suggestions and two anonymous reviewers for their comments on the manuscript.

REFERENCES

Anderson, J. R., Bothell, D., Lebiere, C., & Mantessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language, 38,* 341–380.

Anderton, P. J., & Cole, B. L. (1982). Contour separation and sign legibility. *Australian Road Research, 12,* 103–109.

Buchanan, T. L., Barker, K. N., Gibson, J. T., Jiang, B. C., & Pearson, R. E. (1991). Illumination and errors in dispensing. *American Journal of Hospital Pharmacy, 48,* 2137–2145.

Clark, H. H. (1973). The language-as-fixed-effect fallacy. A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behaviour, 12,* 335–339.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Cohen, M. R. (Ed.). (1999). *Medication errors.* Washington, DC: American Pharmaceutical Association.

Dale, H. C. A., & Baddeley, A. D. (1962). On the nature of alternatives used in testing recognition memory. *Nature, 196,* 93–94.

Davidoff, J. B. (1991). *Cognition through color.* Cambridge, MA: MIT Press.

Davis, N. M., Cohen, M. R., & Teplitsky, B. (1992). Look-alike and sound-alike drug names: The problem and the solution. *Lippincott's Hospital Pharmacy, 27,* 95–110.

Edgar, T. A., Lee, D. S., & Cousins, D. D. (1994). Experience with a national medication error reporting program. *American Journal of Hospital Pharmacy, 51,* 1335–1338.

Flynn, A. F., Barker, K. N., & Carnaham, B. J. (2003). National observational study of prescription dispensing accuracy and safety in 50 pharmacies. *Journal of the American Pharmaceutical Association, 43,* 191–200.

Flynn, E. A., Barker, K. N., Gibson, J. T., Pearson, R. E., Berger, B. A., & Smith, L. A. (1999). Impact of interruptions and distractions on dispensing errors in an ambulatory care pharmacy. *American Journal of Health-System Pharmacy, 56,* 1319–1325.

Flynn, E. A., Barker, K. N., Gibson, J. T., Pearson, R. E., Smith, L. A., & Berger, B. A. (1996). Relationships between ambient sounds and accuracy of pharmacists' prescription-filling performance. *Human Factors, 38,* 614–622.

Hartley, J. (1999). What do these instructions mean? Text design, medical information and older readers. In D. Park, R. Morrel, & K. Shifren (Eds.), *Processing of medical information in aging patients* (pp. 233–248). Mahwah, NJ: Erlbaum.

Hoffman, J. M., & Proulx, S. M. (2003). Medication errors caused by confusable drug names. *Drug Safety, 26,* 445–452.

Johnston, M., McKague, M., & Pratt, C. (2004). Evidence for an automatic orthographic code in the processing of visually novel word forms. *Language and Cognitive Processes, 19,* 273–317.

Kenagy, J. W., & Stein, G. C. (2001). Naming, labeling, and packaging of pharmaceuticals. *American Journal of Health-System Pharmacy, 58,* 3033–3041.

Kohn, L. T., Corrigan, J., & Donaldson, M. S. (2000). *To err is human: Building a safer health system.* Washington, DC: Institute of Medicine.

Lambert, B. L., Chang, K. Y., & Lin, S. J. (2001a). Descriptive analysis of the drug name lexicon. *Drug Information Journal, 35,* 163–172.

Lambert, B. L., Chang, K. Y., & Lin, S. J. (2001b). Effect of orthographic and phonological similarity on false recognition of drug names. *Social Science and Medicine, 52,* 1843–1857.

Lambert, B. L., Donderi, D., & Senders, J. W. (2002). Similarity of drug names: Comparison of objective and subjective measures. *Psychology and Marketing, 19,* 641–661.

Lambert, B. L., Lin, S. J., Chang, K. Y., &, Gandhi, S. K. (1999). Similarity as a risk factor in drug-name confusion errors: The look-alike (orthographic) and sound-alike (phonetic) model. *Medical Care, 37,* 1214–1225.

Reason, J. (1990). *Human error.* Cambridge, UK: Cambridge University Press.

Rogers, W. A., Lawson, N., & Rousseau, G. K. (2000). Warning research: An integrative perspective. *Human Factors, 42,* 102–139.

Silver, N. C., & Braun, C. C. (1993). Perceived readability of warning labels with varied font sizes and styles. *Safety Science, 16,* 615–625.

Smither, J. A., & Braun, C. C. (1994). Readability of prescription drug labels by older and younger adults. *Journal of Clinical Psychology in Medical Settings, 1,* 149–159.

Watanabe, R. K. (1994). The ability of the geriatric population to read labels on over-the-counter medication bottles. *Journal of the American Optometric Association, 65,* 32–37.

Wichmann, F. A., Sharpe, L. T., & Gegenfurtner, K. R. (2002). The contributions of color to recognition memory for natural scenes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 509–520.

Wogalter, M. S., Begley, P. B., Scancorelli, L. F., & Brelsford, J. W. (1997). Effectiveness of elevator service signs: Measurement of perceived understandability, willingness to comply, and behaviour. *Applied Ergonomics, 28,* 181–186.

Wogalter, M. S., & Dingus, T. A. (1999). Methodological techniques for evaluating behavioural intentions and compliance. In M. S. Wogalter, D. M. DeJoy, & K. R. Laughery (Eds.), *Warnings and risk communication* (pp. 53–82). London: Taylor & Francis.

Wogalter, M. S., & Vigilante, W. J. (2003). Effects of label format on knowledge acquisition and perceived readability by younger and older adults. *Ergonomics, 46,* 327–344.

Young, S. L., Laughery, K. R., & Bell, M. (1992). Effects of two type density characteristics on the legibility of print. In *Proceedings of the Human Factors Society 36th Annual Meeting* (pp. 504–508). Santa Monica, CA: Human Factors and Ergonomics Society.

Young, S. L., & Wogalter, M. S. (1990). Comprehension and memory of instructional manual warnings: Conspicuous print and pictorial icons. *Human Factors, 32,* 637–649.

Ruth Filik is a postdoctoral research assistant in the Department of Psychology at the University of Glasgow, UK. She received her Ph.D. in cognitive psychology in 2004 from the University of Derby.

Kevin Purdy is a research associate in the Applied Vision Research Centre at the University of Loughborough, UK. He received his M.Sc. in ergonomics in 1993 from Loughborough University of Technology.

Alastair Gale is head of the Applied Vision Research Centre at the University of Loughborough, UK. He received his Ph.D. in psychology in 1980 from the University of Durham.

David Gerrett is head of pharmacy at the Pharmacy Academic Practice Unit, University of Derby, UK. He received his Ph.D. in pharmacy practice in 1995 from the University of Derby.

Date received: September 25, 2003
Date accepted: November 11, 2004

Downloaded from hfs.sagepub.com at Swansea University on February 16, 2015