A field test of the effects of instruction design on colorectal cancer self-screening accuracy

Markus A. Feufel\(^1\), Tamera R. Schneider\(^1\)* and Hans J. Berkel\(^2\)

\(^1\)Department of Psychology, Wright State University, 3640 Colonel Glenn Highway, Dayton, OH 45435, USA and
\(^2\)Cancer Prevention Institute, 601 West Riverview Avenue, Dayton, OH 45406, USA

*Correspondence to: T. R. Schneider. E-mail: tamara.schneider@wright.edu

Received on January 30, 2009; accepted on February 10, 2010

Abstract

A field experiment tested whether instruction design improves accurate adherence to instructions for medical do-it-yourself tests like the Fecal Occult Blood Test (FOBT). As part of an outreach campaign, 16,073 participants received FOBTs with instructions that were (i) human factored, (ii) motivational, (iii) human factored/motivational combined, or (iv) the standard used in the past. Among all test results returned (\(N = 2483\)), only the human factors instructions reduced errors in filling out result cards. However, after post-validating result cards that had errors, the human-factored, motivational and merged instructions reduced errors. The present findings show that medical instructions designed with human factors and persuasion principles increase accurate adherence. These design principles provide simple and cost-effective ways to increase test taking accuracy and FOBT effectiveness. Better screening instructions can improve the chances of detecting colorectal cancer early, which may help to decrease cancer mortality.

Introduction

The importance of the consumer’s role in their own health promotion and prevention is often underestimated [1]. Patients make the often vital decisions about when and how to engage in preventive care and when to question, comply with, or withdraw from prescribed regimens. One emphasis in research on patient participation in health care is on compliance and adherence generally [2–4] and instruction usability specifically [5–7]. This study expands upon research on instruction usability by complementing human factors design principles with a model of persuasion focusing on motivating health behaviors [8] and by providing a field test of instruction design on the accurate execution of an at-home cancer screening test.

According to the American Cancer Society (ACS), colorectal cancer is the third most common cancer and cause of cancer-related deaths in the US and has been for over 60 years [9]. The most efficient treatment is to remove precancerous tissue and polyps early, before cancer develops. However, during early stages, colorectal cancer is asymptomatic and unlikely to be detected without preventive testing. The ACS recommendations include various tests such as colonoscopies, performed every 10 years, and Fecal Occult Blood Tests (FOBT), every year, for adults over 50 years of age. Yearly FOBTs are efficient [10] and cost-effective for detecting colorectal cancer [11] if supplemented with procedures such as a colonoscopy every 5–10 years. Colonoscopies are performed by gastroenterologists, but FOBTs are performed by patients without supervision or assistance. If patients feel overwhelmed by test instructions or doubt test efficacy, they may not take the FOBT [8]. Recent research found that both human factors and persuasion-derived instructions...
increase test-taking intentions in adults over 50 [12] and increased adherence (i.e. test result return) [T.R. Schneider, M.A. Feufel, H.J. Berkel, ‘The role of FOBT instructions on facilitating adherence to colorectal screening’, in preparation]. Despite increased intentions to be adherent, if FOBT results are inaccurate, the utility of this at-home screening tool is diminished. We examined the influence of human factors and persuasion on FOBT result accuracy.

The main problem for instruction design is clear communication. Instructions are often mismatched to users’ cognitive capacities [5, 13]. Processing barriers may be due to multiple factors such as ill-conceived typographical design [14], weak semantic coherence of text [15] or high-task complexity [16]. Consequently, people may be unable to follow instructions [17] and stop adhering to medical prescriptions [2, 4]. Incorrectly processed health information is uneconomical and may be dangerous.

Research shows that instructions matched to cognitive capacities are preferred and facilitate the acquisition of declarative and procedural knowledge in adults [5, 18, 19]. However, few field studies have examined the effect of instruction design on adherence or its implications for target populations (patients). Further, most studies do not allow conclusions about the effects of instruction design on adherence or its implications for target populations (patients). The present study filled these gaps in a field setting.

Given the step-wise nature of the FOBT, our overall goal in selecting human factors design principles was 2-fold. First, design principles should reduce information to the essential steps needed to perform and record the test. To do so, use-centered design extracts the functional constraints associated with a task so they can be made salient, which should facilitate accurate performance [22]. Second, principles of user-centered design suggest simple wording, icons and a clear layout (boxes or lists) to further reduce cognitive barriers and complexity, which may result in enhanced readability and understanding of semantic relationships [23, 24]. For the target population (adults over 50), minimizing cognitive barriers and clarifying the FOBT procedure should increase comprehension and accurate execution of the instructions [4]. We scrutinized FOBT instructions in use (standard of care) and transformed them by applying human factors/cognitive engineering methods [25]. We focused on content, organization, presentation and language use to convey information (see Table I) [19].

Along with reducing cognitive barriers, optimized instructions should motivate adherence for people who should engage in certain health behaviors [3, 4, 8]. Even if people are aware of their health risks, know about potential remedies and have the opportunity to engage in health behavior, they may hold misconceptions that hamper behavioral engagement. Health messages should provide a valid conspicuous reason for behavioral engagement and clarify procedural effectiveness [8, 26].

The biobehavioral model of persuasion states that effective health communications (i) evoke moderate personal concern about the health threat, which will increase feelings of vulnerability and (ii) enhance efficacy beliefs about health recommendations, both of which trigger behavioral engagement [8]. This model stems partly from Janis’ [26] theory of persuasion, which proposed an inverted U curve to explain the impact of too little or too much

---

**Table I. Summary of human factors guidelines to foster readability, comprehensibility and usability**

<table>
<thead>
<tr>
<th>Use-centered content/organization</th>
<th>User-centered presentation</th>
<th>Language comprehensibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Provide name and purpose</td>
<td>1. Use list not prose format</td>
<td>1. Use simple, clear words</td>
</tr>
<tr>
<td>2. Describe task/procedure</td>
<td>2. Use boxes</td>
<td>2. Use 4–5th grade writing style</td>
</tr>
<tr>
<td>4. List each necessary step</td>
<td>4. Use small discrete sections</td>
<td>4. Use explicit language</td>
</tr>
<tr>
<td></td>
<td>5. Use headings</td>
<td>5. Use positive, active formulations</td>
</tr>
<tr>
<td></td>
<td>6. Increase contrast</td>
<td>6. Use simple icons</td>
</tr>
</tbody>
</table>
emotional arousal on behavioral engagement [26]. A health message that conveys neutral information may trigger little arousal and result in minimal or no reaction. On the other hand, strong arousal elicits defense or avoidance and thwarts behavioral engagement [8, 27, 28]. Compared with messages with highly or minimally threatening content, a message eliciting moderate levels of personal concern should maximize the potential to alter attitudes and intentions [8, 28], partly because they maximize message recipients’ ability and motivation to process the information [29]. Message elaboration should increase comprehension and adherence to instructions [8]. However, moderate personal concern is not enough, efficacy beliefs are ‘the foundation of human motivation and behavior’ (p. 144) and determine whether people engage in or refrain from performing anticipated actions [30]. Health information that reinforces beliefs about the ability (i.e. self-efficacy) or resources (i.e. resource efficacy) to perform actions are better at motivating behavior than messages that solely arouse emotions [31–33]. Threatening information that is not met with adequate levels of self- or resource efficacy beliefs may result in disengagement [34], escape-avoidance behavior [35, 36] and lack of adherence [37]. Disengagement or defensive information processing will interfere with learning how to effectively cope with a health threat and will render elaborate message design useless [8]. For FOBT instructions, providing moderately concerning information about the colorectal cancer threat and information about the benefits and efficiency of the test should challenge individuals to engage in test taking and provide the cognitive processing means to do so accurately.

The data for the present research were provided by the Cancer Prevention Institute (CPI) where health professionals—blind to study manipulations—evaluated the accuracy of FOBT result cards. If errors were detected, phone calls were made to clarify and potentially correct errors, thus post-validating results. If errors could not be corrected, the result card was considered truly invalid and participants were sent a new test. We examined the influence of instruction design on FOBT accuracy by examining the proportion of FOBT result cards returned without errors versus with errors and the proportion of truly valid versus truly invalid result cards. We hypothesized that (i) human-factored or persuasive FOBT instructions would initially result in more accurate result cards compared with the standard and (ii) human-factored or persuasive FOBT instructions would result in more accurate result cards after post-validation compared with the standard. We also expected that the instructions with the combination of human factors and persuasion principles would result in higher accuracy compared with the standard instructions given the initial and post-validated results.

**Methods**

**Framework of the study**

This research was a subproject of a colorectal cancer screening campaign, which was held in March, 2006. The campaign was a collaborative effort of a local TV station, the CPI, a regional ACS and a grocery/pharmacy chain. The goal was to raise public awareness and promote screening of colorectal cancer across four Midwestern metropolitan counties. One month before the campaign, public service announcements were run daily and supported by the TV station’s Website. Large posters were put in stores, near entrances and pharmacies. We did not collect or use personal information beyond that obtained by the campaign; this research was deemed exempted by the internal review board.

**Participants**

A total of 19,683 FOBT kits were distributed to participating pharmacies. After counting test kits not handed out, it was determined that 16,073 test kits were distributed to participants by 27 pharmacies. A total of 2,932 test result cards were returned (17.6%). The return rate is similar to prior campaign years (2004: 15.3% and 2005: 16.5%), which used the same FOBT. Participants returning result cards were mostly female (63.7%) and White (92%), 6.3% of these participants were African-American and 1.5% indicated other ethnicities. The 6.3% African-American participation rate appears to be commensurate with the proportion of African-Americans
spread across the four counties that were participating in the campaign (averaging 9.5%). About 46% participated in the campaign the previous year and 9% did not indicate previous participation. The mean age was 68 years (SD = 11), with a range of 21–93 (6.0% were younger than 50, the recommended age for FOBTs). Most participants were older than 70 (49.2%) followed closely by the 50–70 group (42.4%) and 2.4% did not note age. On average, result cards were returned within 28 days (i.e. during the campaign); however, some were returned up to 128 days.

The FOBT
The FOBT has a specificity, proportion of true negatives, of 98% [38] compared with 99.6% for colonoscopies the most thorough screening procedure [39]. Sensitivities, the proportion of true positives, range between 72 and 78% for FOBTs [38] and 98 and 99% for colonoscopies [39]. The FOBT for this study was a do-it-yourself test with similar specificity and sensitivity values (Biomerica, Newport Beach, CA, USA) [40]. Each test included five blood-sensitive tissues, a pouch with control powder to check the functioning of the tissues, a result card and instructions. After finalizing the test procedure, instructions requested users to mail the pre-addressed stamped result card for professional evaluation (addressed to CPI).

Result card validation
Adherence was denoted as result card return. Upon initial receipt, result cards were determined by a CPI health professional with expertise in cancer screening to be filled out accurately or inaccurately. Some result cards could be post-validated in that errors were rectified by contacting the user (e.g. forgetting to write down a date of a reading) or they were truly invalid (e.g. skipping one of the five tests, leaving too much time between test dates).

If the test results indicated blood in the stool, participants were notified and follow-up examinations recommended. One and a half months after the closing date of the campaign (24 March 2006), the CPI provided us with the numbers of correctly filled out result cards, post-validated cards and truly invalid cards. In addition, CPI provided a database with other information written on the result cards, including the result card identifier number, date of result card return, participation in last year’s campaign and demographics (ethnicity, age and sex).

Stimuli
To study the influence of instruction design on FOBT accuracy, the instructions were varied systematically yielding a two-by-two factorial design (human factors: low and high; motivation: low and high). The standard instructions hitherto provided with the FOBT were used for crafting the other three instruction sets.

Human-factored instructions
Figure 1 provides a miniaturized example of the standard (top) and human-factored (bottom) instructions. Table I shows four use-centered features extracted from the content provided in the standard instructions [2, 6]. These features were arranged in user-preferred order [7] and presented as they relate to knowledge and objectives [41]. The purpose of the FOBT was presented, followed by a description of the procedure, possible outcomes and side effects [42, 43]. The major steps required were extracted and used to prime users for the upcoming task [25]. That is, an introductory section specified the ‘specific procedure’ for testing for blood in the stool, how to ‘distinguish’ a positive from a negative result and how to ‘record’ and ‘report’ results. After establishing the FOBT procedural context, detailed directions for each step were provided on a separate page. Each step of the screening procedure was scrutinized from a use-centered perspective by creating a flow chart of the essential steps involved in each of the five readings required. The human-factored version had a word count of 523, which was significantly lower than the 937 words used in the standard instructions ($\chi^2(1) = 135.36, P < 0.01$).

Table I, column two, summarizes user-centered design principles that were applied to improve the presentation of information. First, list format was maintained and reinforced by using boxes to
Designing cancer screening instructions to increase accuracy

Fig. 1. Illustration of standard (top) and human-factored instructions (bottom).
segregate important facts [23, 24]. Capitalization and short discrete sections with boldfaced headings helped to emphasize importance [25]. To increase readability, contrast was augmented using a high-quality printer [5]. Where possible, written instructions were replaced with or supported by simple icons [44] eliminating the need for orthographic decoding [2, 45, 46].

Finally, the last column of Table I displays guidelines pertaining to language use in instruction design. Language comprehensibility was increased by using unambiguous simple wording, a descriptive 4th–5th grade level writing style, and argument overlap [2]. Implicit information was restated explicitly [15] and negatively framed passive sentences were reformulated as positively framed active sentences [2].

Motivational instructions
The standard FOBT instructions were used as the starting point for the application of the biobehavioral model of persuasion [8], which guided the motivational manipulation. Table II shows that information concerning colorectal cancer was reformulated to evoke moderate levels of ‘personal concern’ about colorectal cancer by providing a rationale for taking the FOBT. Herein, factual information concerning the colorectal health threat was included at the beginning of instructions (e.g. ‘One out of 18 adults will be diagnosed with colon cancer in their lifetime’ and ‘Colon cancer grows without symptoms you can see or feel.’). Sentences aimed at increasing self- and resource efficacy were included in both the introductory and the closing sections of the instructions. Research indicates that people lack confidence in the FOBT procedure and their ability to perform the ‘messy’ procedure [38, 49], so it was presented as a simple and efficacious response to the colorectal health threat (self-efficacy: ‘The test is simple and does not require stool handling’; resource efficacy: ‘This test can help you find colon cancer early, before symptoms develop’). The focus for this instruction set was to change emphasis in wording rather than factors associated with reading ease. The word count increased slightly but non-significantly from 937 for the standard to 1004 for the motivational instructions ($\chi^2(1) = 0.49$, $P > 0.05$), the RES for the motivational version was similar to the standard instructions and both were at an 8th grade reading level.

Merged instructions
To create the merged version, sentences from the motivational version were added to the introductory and closing sections of the human-factored version. The word count increased slightly but non-significantly from 523 for the human factored to 581 for the merged version ($\chi^2(1) = 3.05$, ns); the RES of 78 and 5th grade reading level were similar to the human-factored version (RES of 79).

All four instruction sets were identical with regard to the order in which information was presented. Finally, manipulation tests conducted in a previous study confirmed the effectiveness of both human factors and motivational principles for the target population, adults over 50 years of age [12].

Procedure
FOBTs were provided in a large envelope with one of the four instruction versions and a result card. To identify instruction version, an identifier number on the result card was matched to instruction
Designing cancer screening instructions to increase accuracy

set (e.g. 1–5000 were matched to the standard, 5001–10 000 were with the human-factored version). To assure randomization, four piles of 25 envelopes of each instruction set were shuffled numerous times and placed in boxes of 100. These boxes were sent to pharmacies, where pharmacists distributed them on a first-come, first-served basis. Neither the volunteers who stuffed the envelopes (with FOBTs, result cards and instructions) nor employees of the distributing pharmacies, nor the health professional who evaluated returned result cards was informed about the study or the experimental manipulation.

**Results**

The overall result card return rate, by instruction set, was 16.5% for the standard, 19% for the human-factored, 17.6% for the motivational and 17.5% for the merged instructions. To examine the effect of instruction design on accurate adherence, our sample of 2932 returned test result cards was divided into two groups: initially valid (n = 2151 or 73% of all returned result cards) and initially invalid (n = 781). Splitting the sample according to the instruction version used, Table III shows that initially, there were fewer result cards with errors for the human-factored instructions (24.4%), followed by the merged (25.1%) and motivational versions (26.8%). The standard instructions resulted in the highest error rate (30.7%). After post-validation, a similar pattern emerged for the truly valid [n = 2151 + 250 (post-validated) = 2401 or 82% of all returned result cards] and truly invalid result cards (those that could not be validated by contacting the sender; n = 531 or 18%). There were fewer truly invalid result cards for the human-factored

| Table II. Examples of motivational manipulations compared with standard instructions |
|--------------------------------------|--------------------------------------------------------------------------------------------------|
| Information on colorectal cancer in standard instructions | In the early stages of colorectal disease, many people do not feel sick. Screening tests are used to try to find disease in this early stage. |
| Enhancement of personal relevance | The facts: one of 18 adults will be diagnosed with colorectal cancer in their lifetime. Unfortunately, in the beginning, colorectal cancer grows without symptoms you can see or feel. After symptoms occur, treatment is very difficult. |
| Test purpose and functioning in standard instructions | Blood hidden in your stool (or bowel movement) may show that you have colorectal disease. The EZ-Detect test is to see if you have blood hidden in your stool. The test uses tissues that you float in the toilet with a bowel movement to show hidden blood. |
| Enhancement of resource efficacy | There is good news: this screening test can help you find colorectal cancer early, when treatment is most successful. Blood in your stool may be an early sign of colorectal cancer. This test will find if blood is present in your stool. |
| Enhancement of self-efficacy | The test is simple and does NOT require stool handling. After a bowel movement, you float a test tissue in the toilet. If the tissue changes color, blood is present in your stool. |

| Table III. Total numbers and percentages of returned, initially invalid, post-validated and truly invalid result cards, by instruction set |
|--------------------------------------|--------------------------------------------------------------------------------------------------|
| Result cards | Results cards, by instruction set | Total |
| | Standard | Human factored | Motivational | Merged | |
| Returned | 691 | 780 | 740 | 721 | 2932 |
| Initially invalid | 212 | 190 | 198 | 181 | 781 |
| Initially invalid/returned | 30.7% | 24.4% | 26.8% | 25.1% | 26.6% |
| Post-validated | 58 | 65 | 68 | 59 | 250 |
| Truly invalid | 154 | 125 | 130 | 122 | 531 |
| Truly invalid/returned | 22.3% | 16.0% | 17.6% | 16.9% | 18.1% |
instructions (16%), followed by the merged (16.9%) and motivational (17.6%) instructions. The standard instructions resulted in the highest rate of truly invalid result cards (22.3%).

Data analysis

Two sets of analyses examined the effects of instruction design on (i) initial valid versus initial invalid result cards and (ii) valid (initial and post-validated) versus truly invalid result cards. For each analysis, a three-step hierarchical logistic regression analysis was computed to compare the proportion of cases in each category (valid versus invalid result cards) to the proportion we might expect in each category if there were no effect of instruction design (the null hypothesis). In the first step, covariates were controlled for demographics (e.g. age, sex, ethnicity), past participation (a proxy for experience) and result card return date (a proxy for conscientiousness). The second step examined the main effects of the human factors and motivational manipulations on return rate. In the final step, the interaction of the experimental manipulations was entered. In addition, a planned contrast tested whether the merged instructions improved accuracy compared with the other versions. The results of the logistic regression analyses are summarized in Table IV.

Analysis 1: comparing initially valid to initially invalid result cards

To analyze the effect of instruction design on overall test taking accuracy, a dichotomous variable was created denoting a valid (0) or invalid (1) result card. This outcome was subjected to a three-step hierarchical logistic regression analysis with the total returned result cards (n = 2483) as the base rate.

Table IV. Hierarchical logistic regression analysis of instruction design on result card accuracy

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Analysis 1 (N = 2483)</th>
<th>Analysis 2 (N = 2483)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b</td>
<td>SD</td>
</tr>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>0.30</td>
<td>0.10</td>
</tr>
<tr>
<td>Age</td>
<td>0.03</td>
<td>0.01</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>0.23</td>
<td>0.10</td>
</tr>
<tr>
<td>Mail date</td>
<td>0.01</td>
<td>0.003</td>
</tr>
<tr>
<td>Prior participation</td>
<td>0.19</td>
<td>0.10</td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>0.30</td>
<td>0.10</td>
</tr>
<tr>
<td>Age</td>
<td>0.03</td>
<td>0.01</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>0.23</td>
<td>0.09</td>
</tr>
<tr>
<td>Mail date</td>
<td>0.01</td>
<td>0.003</td>
</tr>
<tr>
<td>Prior participation</td>
<td>0.20</td>
<td>0.10</td>
</tr>
<tr>
<td>HF</td>
<td>-0.29</td>
<td>0.10</td>
</tr>
<tr>
<td>Motivation</td>
<td>-0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>Step 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>0.30</td>
<td>0.10</td>
</tr>
<tr>
<td>Age</td>
<td>0.03</td>
<td>0.01</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>0.23</td>
<td>0.09</td>
</tr>
<tr>
<td>Mail date</td>
<td>0.01</td>
<td>0.003</td>
</tr>
<tr>
<td>Prior participation</td>
<td>0.20</td>
<td>0.10</td>
</tr>
<tr>
<td>HF</td>
<td>-0.72</td>
<td>0.30</td>
</tr>
<tr>
<td>Motivation</td>
<td>-0.52</td>
<td>0.30</td>
</tr>
<tr>
<td>HF X motivation</td>
<td>0.29</td>
<td>0.19</td>
</tr>
<tr>
<td>Planned contrast</td>
<td>-0.13</td>
<td>0.11</td>
</tr>
</tbody>
</table>

OR, odds ratio; CI, confidence interval; HF, human factors. *P < 0.05; **P < 0.01.
This number differs from the overall number of returned result cards \((N = 2932)\) due to missing demographic data \((n = 449)\), which reduced the number of result cards available for the regression analyses.

In the first step, the outcome variable \((1848\text{ valid versus } 635\text{ invalid result cards}; n = 2483)\) was regressed on five covariates (i.e. sex, age, ethnicity, mail date and prior experience). The model was significant, \(\chi^2(5) = 69.80, P < 0.01\). Participants who were female \((\text{Wald}(1) = 8.87, P < 0.01)\), older \((\text{Wald}(1) = 42.39, P < 0.01)\), ethnic minorities \((\text{Wald}(1) = 6.82, P < 0.01)\), mailed in their results later \((\text{Wald}(1) = 13.58, P < 0.01)\) and prior participants \((\text{Wald}(1) = 3.87, P < 0.05)\) were more likely to send in result cards with errors compared with male, younger, White, quicker responding and inexperienced participants.

In the second step, the main effects of human factors and motivation were entered. The model \(\chi^2\) revealed significant improvement over the covariate model, \(\chi^2(2) = 10.81, P < 0.01\). Controlling for the covariates, the motivational manipulation was not significant \((\text{Wald}(1) = 2.17, ns)\), but the human-factored instructions led to fewer truly invalid results \((\text{Wald}(1) = 8.96, P < 0.01, \text{odds ratio} = 0.72)\). In the last step, the interaction was entered. The model fit increased significantly compared with Step 2 \((\chi^2(1) = 3.70, P = 0.05)\), the beta weight of the interaction was significant \((\text{Wald}(1) = 3.70, P = 0.05, \text{odds ratio} = 1.53)\), and the main effects for both human factors and motivation were significant. The planned contrast (merged versus all others) was non-significant \((\text{Wald}(1) = 0.96, ns)\). However, post-hoc comparisons of the differences between the conditions revealed that the human-factored \((\text{Wald}(1) = 12.03, P < 0.01; \text{odds ratio} = 0.58)\), motivational \((\text{Wald}(1) = 5.63, P < 0.05; \text{odds ratio} = 0.7)\) and merged instructions \((\text{Wald}(1) = 9.03, P < 0.01; \text{odds ratio} = 0.63)\) significantly decreased the likelihood of returning truly invalid result cards compared with the standard instructions.

**Analysis 2: comparing truly valid to truly invalid result cards**

Analysis 2 examined the effects of instruction design on the proportion of valid (initial valid and post-validated; \(n = 2059\)) to truly invalid \((n = 424)\) result cards. In the first step, the outcome was regressed on the five covariates. The model was significant, \(\chi^2(5) = 39.33, P < 0.01\). Older \((\text{Wald}(1) = 25.46, P < 0.01)\) and minority participants \((\text{Wald}(1) = 10.53, P < 0.01)\) were more likely to return truly invalid results compared with younger and White participants. Neither sex \((\text{Wald}(1) = 3.69, ns)\), return date \((\text{Wald}(1) = 0.56, ns)\) nor prior participation \((\text{Wald}(1) = 1.75, ns)\) predicted accuracy.

Main effects were entered next. The \(\chi^2\) revealed significant improvement over the covariate model, \(\chi^2(2) = 10.81, P < 0.01\). Controlling for the covariates, the motivational manipulation was not significant \((\text{Wald}(1) = 2.17, ns)\), but the human-factored instructions led to fewer truly invalid results \((\text{Wald}(1) = 8.96, P < 0.01, \text{odds ratio} = 0.72)\). In the last step, the interaction was entered. The model fit increased significantly compared with Step 2 \((\chi^2(1) = 3.70, P = 0.05)\), the beta weight of the interaction was significant \((\text{Wald}(1) = 3.70, P = 0.05, \text{odds ratio} = 1.53)\), and the main effects for both human factors and motivation were significant. The planned contrast (merged versus all others) was non-significant \((\text{Wald}(1) = 0.96, ns)\). However, post-hoc comparisons of the differences between the conditions revealed that the human-factored \((\text{Wald}(1) = 12.03, P < 0.01; \text{odds ratio} = 0.58)\), motivational \((\text{Wald}(1) = 5.63, P < 0.05; \text{odds ratio} = 0.7)\) and merged instructions \((\text{Wald}(1) = 9.03, P < 0.01; \text{odds ratio} = 0.63)\) significantly decreased the likelihood of returning truly invalid result cards compared with the standard instructions.

**Discussion**

The present study investigated the effects of human factors and motivational instruction design on the accuracy of filling out test result cards for a do-it-yourself colorectal cancer screening test. In Analysis 1, we compared the proportion of errors in the initial group of result cards received and found that the human factors manipulation decreased errors, as expected, whereas the motivational manipulation did not. In terms of the absolute reduction of invalid result cards, the human-factored and merged instructions compared with the motivational and standard sets resulted in about four fewer invalid result cards for every 100 that have been returned. (Together, the human-factored and merged instruction sets resulted
in 371 of 1501 or about 25 invalid result cards for every 100 that have been returned. The motivational and standard instruction sets, on the other hand, resulted in 410 of 1431 or about 29 invalids of every 100 returned result cards. Thus, in absolute numbers, the human factors manipulations resulted in about four fewer invalid result cards for every 100 that have been returned.) For the 2932 returned result cards, this implies 117 fewer cases requiring post-validation.

One explanation that the human factors but not the motivational manipulation reduced result card errors is that there was a more comprehensive approach taken for the human-factored set. The human factors manipulations focused on content, organization and layout (i.e. presentation and language use). The motivational manipulation focused on content alone, increasing personal relevance and efficacy beliefs. The motivational information was launched from the standard instruction base, yielding similar low readability ease scores and high word counts, compared with the human-factored versions. Such inherited features may have reduced the effectiveness of the motivational manipulations so that their effect could only be observed after post-validating initially invalid result cards. Perhaps, the effectiveness of motivational manipulations also depends on how simply information is presented. However, the relative ineffectiveness of the merged instructions for reducing errors in the initially mailed-back result cards speaks to a different kind of interactive relationship between information presentation and motivational effects. This idea will be further explored after summarizing Analysis 2. We also found that other variables influenced result card accuracy in that men, those younger, Whites, those returning their results early and those with prior participation were more accurate than their counterparts, which will also be explored further below.

Analysis 2 investigated the effect of instruction design on the proportion of result cards still invalid after post-validation versus all valid result cards. There was a significant interaction of human factors and motivational manipulations on accurate result card return. Post-hoc comparison showed that for this group of result cards, all instructions with manipulations significantly reduced the proportion of invalid results. Specifically, compared with the standard instructions, the human-factored instructions resulted in about 6, the motivational version in about 4 and the merged instruction set in about 5 fewer invalid result cards for every 100 that have been returned. (The standard instruction set resulted in 154 of 691 or about 22 invalid result cards for every 100 that have been returned. The human-factored version resulted in 125 invalids of 780 returned result cards (16 of 100), the motivational set in 130 invalids of 740 returned results card (18 of 100) and the merged instruction version in 122 invalids of 721 returned results card (17 of 100). Thus, compared with the standard instruction set, the absolute reduction of invalid result cards for the human-factored version is about 6 of 100 returned result cards, 4 of 100 for the motivational and 5 of 100 for the merged instruction set.) The merged instructions did not reduce result card errors compared with either the human-factored or the motivational instructions. With respect to the 2932 returned result cards, these findings imply an additional 117–176 correct cancer screening results. This suggests that both manipulations effectively reduced result card errors, which invalidated FOBT results and hampered screening efforts.

Given that human factors principles had an effect on the accuracy of initial and total validated result cards, it seems likely that the motivational instructions, though effectively decreasing the proportion of invalid result cards in the second analysis, did not add incrementally to the effectiveness of the human-factored instructions. One explanation may be that human factors improvements increased user’s beliefs about their ability to take the test accurately (i.e. self-efficacy). Indeed, preliminary studies found that adults over 50 rated visuals and graphic elements as helpful for structuring and understanding the test procedure [12]. Whereas the motivational manipulations may have increased personal relevance and resource–efficacy [8], the human factors improvements may have increased participants’ perceived ability to follow the procedure (i.e. self-efficacy and message elaboration). That is, human factors principles and motivational manipulations may both work to increase efficacy beliefs.
Indeed, there is evidence that text features such as the ones induced by the human factors manipulations may increase motivation to process information [47]. Flesch [47] suggested a ‘human interest’ score (This score is a function of the proportion of personal pronouns per total words and the proportion of ‘personal sentences’ addressing the reader compared with the total sentences. Values close to 100 denote more interest.), a measure of appeal, to complement scores measuring reading ease. Interestingly, both the human-factored and the merged instructions yielded scores categorized as ‘highly interesting’ with values of 53 and 45, respectively. However, the motivational instructions bordered between highly interesting and ‘interesting’, with a score of 40, and the standard version were in the range of interesting with a score of 35. Whereas these scores suggest that human factors principles increased the appeal of the instructions, our prior research found that human-factored changes are perceived as helpful to executing the instructions [12]. Both factors may have worked to outweigh or mingle with the effects intended by the motivational manipulations based on the biobehavioral model of persuasion (personal relevance and efficacy). Given that both human factors (Analysis 1 and 2) and motivational principles (analysis 2) effectively reduce errors in result cards, there may be an upper limit to the impact instruction design that can have on users’ motivation. Future research in a more controlled setting could test this assertion.

Our lack of consistent findings for the motivational instruction set could be also related to various factors associated with this sample. First, information can only be motivational if it is understood and personally meaningful. In the present research, we took a previously existing set of instructions and systematically varied them. For the persuasive instruction set, the existing instructions were only slightly modified. Given the verbose and dense standard instructions, the persuasive modifications were set up to be an extremely conservative test of persuasion. In addition, the persuasive addition itself may have been more effective if it was worded a little differently. In an African-American sample, specific information about their group was preferred to a general statement about cancer statistics, such as was used in the present study (e.g. ‘1 in 18’) [50]. However, it is not clear that this preference would have influenced behavior in this mostly White sample who may implicitly assume that the statistics refer to themselves. Still, it would be interesting to examine whether a statement such as ‘think of 18 family members and friends, one of them will develop colorectal cancer in their lifetime’ would have had more impact [50]. Another reason that the motivational manipulations may not have been as helpful in facilitating accurate result card return is that, compared with non-participants, those who participated may have higher risk perceptions associated with a family history, a personal history of polyps or worry more about developing colorectal cancer [51]. This suggests that public health outreach campaigns should not only evaluate the effectiveness of their programs but also strive to understand their participants and target populations across multiple levels, perhaps including investigation of individual-level factors that might influence campaign participation. Only through program, evaluation will campaigns grow to reach more participants, more effectively.

Among our covariates, older adults were more likely to return initially and truly invalid result cards. The ability to read, understand and remember printed or spoken instructions tends to decline among those over 50. This may be due to decreased sensory functioning, working and procedural memory and processing speed. This decline requires instruction strategies that explicitly inform older users how to adhere and why it is important [6]. As our population ages, research on the eldest of the old is warranted. There may be cognitive and capability limitations in this age group of which we are yet unaware. Those who were not White were more likely to return initially invalid and truly invalid result cards. This is an area of inquiry in the human factors domain that is virtually non-existent, but is warranted. The lack of accuracy in the non-White groups is particularly concerning as outreach campaigns are particularly concerning as outreach campaigns are often geared toward increasing participation among the medically underserved, which includes ethnic minorities. It may be that appropriately targeted communications could enhance...
motivation, information processing and accurate screening in this group [52].

In sum, the present findings suggest that human factors and motivational considerations are helpful in instruction design to increase accuracy in returning valid result cards. However, their effect on accurate use may not be as simple as reflected in the present research design. Instead, the above findings, supplemented by past research, suggest that human factors improvements may simultaneously influence readability and appeal. On the other hand, the effectiveness of motivational instruction design may also depend on how motivational information is presented or ‘human factored.’ Whereas the present study suggests that the effect of the human factors manipulations was most robust, future research is needed to disentangle the interactive effect of human factors and motivational message design on appeal, personal relevance, efficacy beliefs and ultimately the accurate use of instructions.

Limitations

This experiment has several limitations related to participant population, the ecological setting which provided the framework for this field study and the nature of the manipulations. First, participants were predominantly White (92%) and middle class. This may have partly occurred because the FOBT kits were distributed across counties serving mostly these populations. Future research is needed to investigate whether the applied design principles may complement socioeconomic and culturally targeted strategies to promote colorectal cancer prevention to minorities and at-risk populations [53]. Second, a test kit was given to anyone, but only motivated participants may have requested it, which may have diminished the effect of the motivational manipulation on adherence. In addition to this self-selection bias, our operationalization of the dependent variable was limited to the accuracy of returned result cards, but we had no information about the underlying affective, cognitive and behavioral processes that may have led to errors. For example, errors may reflect frustration, limited understanding or problems remembering tests outcomes. Cognitive task analysis techniques may provide a powerful method to investigate how instructions are used in homes, identify the problems users face during the procedure [54] and help implement improved instruction designs. Lastly, the applied manipulations did not allow meaningful conclusions about what made instructions more effective. Reduced procedural complexity, improved readability or the whole ‘gestalt’ of the redesign may have mattered most. Moreover, how these aspects interacted with the motivational manipulation for the effective merged instruction set is unclear. Systematic follow-up should investigate these important questions.

Theoretical implications

Several researchers have suggested that apart from financial, educational and psychosocial measures, the application of human factors principles may play an important role in increasing patient adherence [4, 19, 55, 56]. Many studies on human-factored instructions report that they are preferred [2, 5, 24], and easier to understand [57, 58] and memorize [7, 43], compared with instructions that lack human factors considerations. To date, most intervention studies that employ human factors principles do not yield meaningful conclusions because the effect of medical instructions are confounded with other aspects of interventions (e.g. personal tutoring, follow-up) [20, 21]. Thus, the present research is, to our knowledge, one of the first systematic field studies to show a positive relationship between human-factored instructions and accurate adherence in users. Studies concerning motivation and message framing have focused on motivating actual adherence. For example, research focused on the differential effect of loss versus gain framed health information on smoking cessation [59], Pap testing [60] or mammography screening in specific ethnic and/or socioeconomic groups [52]. Research using the biobehavioral model of persuasion has motivated detection behaviors [8] and increased FOBT intentions [12]. The present study extends past research by showing a decreased likelihood of irrecoverable mistakes with motivational or merged instructions (Analysis 2). The motivational manipulations likely facilitated personal concern and higher efficacy, which may have fostered more thorough
message processing compared with the standard instructions [8]. As for the human factors manipulation, future studies will need to explore the pathways of how the applied motivational principles impact accurate adherence to medical instructions.

**Practical implications**

As stated in a recent New York Times article on ‘How Cancer Rose to the Top of the Charts’ [61], there is no panacea against cancer. The most promising remedy is hitherto prevention and early detection of precancerous cells. FOBT adherence can be increased by providing pre-stamped result cards [62], mailing FOBTs [63] or providing incentives [64]. Similarly, reducing cognitive barriers increases FOBT return rates [T.R. Schneider, M.A. Feufel, H.J. Berkel, ‘The role of FOBT instructions on facilitating adherence to colorectal screening’, in preparation]. The present research adds to these findings by showing that human factors and motivational instruction features facilitate result card accuracy and, compared with standard instructions, may decrease the number of invalid result cards by four to six of every 100 result cards that are returned. Human-factored instructions reduced the health care workload because fewer results required post-validation. Further, once validated, instructions that had human-factored and motivational elements helped patients and health care professionals to obtain more accurate cancer screening results. Applied en masse, the cognitive and motivational optimization of FOBT instructions may be one more step toward decreasing cancer mortality rates.

**Funding**

Ohio Board of Regents.

**Conflict of interest statement**

None declared.

**References**

19. Morrow DG, Weiner M, Deer M *et al*. Patient-centered instructions for medications prescribed for the treatment...
58. Morrell RW, Park DC, Poon LW. Quality of instructions of prescription drug labels: effects on memory and


