Patient Education and Counseling 92 (2013) 260–265

Contents lists available at SciVerse ScienceDirect

Patient Education and Counseling

journal homepage: www.elsevier.com/locate/pateducou

Intervention

The effects of two health information texts on patient recognition memory: A randomized controlled trial

Erin Freed [a], Debra Long [a], Tonantzin Rodriguez [b], Peter Franks [c], Richard L. Kravitz [d], Anthony Jerant [c,*]

[a] Department of Psychology, University of California Davis, Davis, USA
[b] University of California Davis School of Medicine, Sacramento, USA
[c] Department of Family and Community Medicine, University of California Davis School of Medicine, Sacramento, USA
[d] Department of Internal Medicine, University of California Davis School of Medicine, Sacramento, USA

ARTICLE INFO

Article history:
Received 8 October 2012
Received in revised form 27 February 2013
Accepted 9 March 2013

Keywords:
Colorectal neoplasms
Early detection of cancer
Educational status
Health literacy
Mass screening
Mental recall
Patient education as topic
Patient education handout
Randomized controlled trial
Reading

ABSTRACT

Objective: To compare the effects of two health information texts on patient recognition memory, a key aspect of comprehension.
Methods: Randomized controlled trial ($N = 60$), comparing the effects of experimental and control colorectal cancer (CRC) screening texts on recognition memory, measured using a statement recognition test, accounting for response bias (score range −0.91 to 5.34). The experimental text had a lower Flesch–Kincaid reading grade level (7.4 versus 9.6), was more focused on addressing screening barriers, and employed more comparative tables than the control text.
Results: Recognition memory was higher in the experimental group (2.54 versus 1.09, $t = -3.63$, $P = 0.001$), including after adjustment for age, education, and health literacy ($\beta = 0.42$, 95% CI: 0.17, 0.68, $P = 0.001$), and in analyses limited to persons with college degrees ($\beta = 0.52$, 95% CI: 0.18, 0.86, $P = 0.004$) or no self-reported health literacy problems ($\beta = 0.39$, 95% CI: 0.07, 0.71, $P = 0.02$).
Conclusion: An experimental CRC screening text improved recognition memory, including among patients with high education and self-assessed health literacy.
Practice implications: CRC screening texts comparable to our experimental text may be warranted for all screening-eligible patients, if such texts improve screening uptake.
© 2013 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Next to verbal interactions in face-to-face or telephonic patient-provider encounters, text documents are the most common way of providing patient health information [1,2]. Given the high prevalence of low literacy in the United States (U.S.) [3], a general recommendation for maximizing patient comprehension is to make health information texts as simple as possible (e.g. by reducing reading level), without losing key context and meaning [1]. While observational studies provide support for this recommendation [4], evidence from randomized controlled trials (RCTs) is scant.

In the past 30 years, only seven published RCTs have explored the effects of simplified health information texts on comprehension, with mixed findings [5–11]. Focus health issues in the various RCTs were warfarin use [7], human immunodeficiency virus infection risk [11], polio vaccination [5,6], smoking [9], and informed consent for experimental chemotherapy [8,10]. The limited number of trials in this realm, with differing health topics and mixed findings, suggest the need for further RCTs comparing comprehension of different health information texts, encompassing additional health topics.

No RCTs have compared patient comprehension of colorectal cancer (CRC) screening documents that differ in design and content focus. This is a key research gap given that CRC screening is a relatively complex health topic given several available test options, each with differing pros and cons [12]. Perhaps, in part, for this reason, CRC screening knowledge and uptake are low in the U.S. population relative to other evidence-based cancer screening tests [13–16]. Further, no trials have addressed whether patient education level and health literacy may influence comprehension of different CRC screening texts. This is also important to research, since CRC screening knowledge and uptake are low among less educated and less literate persons in the U.S. [13–16]. Theoretically, text design features anticipated to facilitate comprehension (e.g. lower reading level, use of comparative tables) should most benefit persons with low education and literacy. Of particular clinical interest and practical importance is whether patient *self-assessed*

* Corresponding author at: Department of Family and Community Medicine, University of California Davis School of Medicine, 4860 Y Street, Suite 2300, Sacramento, CA 95817, USA. Tel.: +1 916 734 7081; fax: +1 916 734 5641.
E-mail address: afjerant@ucdavis.edu (A. Jerant).

0738-3991/$ – see front matter © 2013 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.pec.2013.03.008

health literacy is associated with comprehension of different texts. National quality improvement blueprints encourage universal health literacy assessment in clinical settings [1], yet time constraints often preclude the use of objective measures, prompting exploration of employing brief patient self-assessment "screeners" [17–19].

Prior RCTs comparing patient comprehension of different health texts used various measures to assess comprehension, a complex, multi-faceted construct (process) that cannot be directly observed, and for which controversy exists regarding optimal measurement [20]. None of the prior RCTs sought to measure *inferential comprehension*, or inference of meanings not directly explained (i.e. implicit) in a text, a conceptually high level activity of clear relevance to health education and behavior [21]. However, validly and reliably measuring inferential comprehension is difficult, since it is strongly dependent on reasoning skills [22]. Fortunately, assessing inferential comprehension may not be critical when comparing the degree to which different texts *explicitly* convey basic information regarding a given health topic, since this is essentially a matter of lower level or *literal comprehension* [21]. Furthermore, literal comprehension is a prerequisite to and predictive of inferential comprehension [21–24]. Likely for these reasons, prior RCTs have examined text effects on one or both of two aspects of literal comprehension: *recall memory*, the ability to remember elements of a previously viewed text without visual prompting [5,6,8,11]; and *recognition memory*, the ability to accurately recognize previously viewed information when encountered again in written form [7,9–11].

The recall memory measures in prior RCTs employed open-ended verbal questioning of participants, requiring study personnel judgment in determining the appropriateness of responses, potentially resulting in bias. The recognition memory measures in the prior RCTs were written multiple choice and/or true-false items, which are vulnerable to educated guessing and response bias (e.g. a tendency to prefer true to false answers) [25]. Verbal open-ended recall memory questions and written multiple choice and true-false recognition memory questions also are susceptible to confounding by pre-existing knowledge of the health topic, since they typically do not require participants to correctly identify *verbatim* passages (e.g. complete sentences) from viewed texts.

Employing a signal detection theory-grounded approach to measuring recognition memory can help to minimize the effects of response bias and background knowledge confounding on recognition memory scores, providing a purer estimation of the effects of texts themselves on literal comprehension. Signal detection theory recognizes that most human decisions are made under conditions of uncertainty [26]. The theory further recognizes that under such conditions, human judgments do not always arise from a fully balanced, well-reasoned, and accurate assessment of the situation, but instead are often driven largely or fully by educated guessing, innate biases (e.g. response option preferences), or the overriding influence of background contextual knowledge. These underlying tenets of signal detection theory have been employed to inform an approach to measuring recognition memory that minimizes response bias and background knowledge confounding. Briefly, a written recognition memory test is developed incorporating an equal number of verbatim statements extracted from each study text being compared in a RCT [27,28]. Study participants are then asked to identify the statements that appeared in their randomly assigned text. Both correctly identified statements ("hits" – a measure of sensitivity) and incorrectly identified statements ("false alarms" – those that had actually appeared in the other study text – to capture response bias effects) are employed to calculate a summary discriminability or *d* prime (d') score – essentially, an indicator of the "true signal" relative to "noise" (bias and confounding effects) in participant responses. This approach is well-established in psycholinguistic and cognitive science studies but, to our knowledge, has not been used in text comprehension RCTs in the biomedical realm [27–29].

We conducted a RCT, comparing patient recognition memory of an experimental colorectal cancer screening (CRC) information text and of a control CRC screening text. The experimental text was written at a lower Flesch–Kincaid reading grade level, focused more on addressing practical CRC screening barriers, and relied more heavily on tabular presentation of information than the control text. We also explored the roles of patient education level and self-assessed health literacy in influencing text recognition memory. We employed a written signal detection theory-grounded measure to derive a recognition memory d' score, accounting for both item recognition sensitivity and response bias effects. We hypothesized that: (1) compared with controls, experimental group patients would have better recognition memory of their randomly assigned text; and (2) the benefit in recognition memory would be restricted to patients with less education and lower self-assessed health literacy.

2. Methods

2.1. Study setting, sample recruitment, and randomization

Study activities were conducted from September 2009 through March 2010. The local institutional review board approved the study (ClinicalTrials.gov identifier: NCT00965965).

English-speaking persons aged 50–75 years receiving primary care from a family physician or general internist at one of two offices in the Sacramento, California area were telephoned to solicit their participation. The lower and upper age cut points for study participation were selected based on U.S. Preventive Services Task Force evidence-based CRC screening guidelines, which recommend routine screening in all adults aged 50–75 [12]. Patients were asked whether they had received fecal occult blood testing (FOBT) within the past year, flexible sigmoidoscopy within 5 years, or colonoscopy within 10 years. Those answering "no" to these questions and reporting adequate eyesight to read printed text were eligible to participate.

Eligible patients who agreed to participate met with study personnel at a central location, where written informed consent was obtained, followed by random assignment to read one of the study texts. Randomization was at the level of the individual patient, implemented in blocks of 10 patients to help ensure a balance in sample size across the two study groups over time [30], using sealed shuffled envelopes containing group assignments. We estimated that a sample of 44 patients (22 per group) would yield 90% power to detect a small effect (0.3 standard deviations) on recognition memory, the outcome of interest. We conservatively targeted recruitment of 60 patients (30 per group), to ensure an adequate sample in the event of attrition or missing data.

After randomization, participants completed a pre-intervention questionnaire, read their assigned text, and then completed a comprehension test (see Section 2.3). Participants received a $30 gift card after completing these activities.

2.2. Study texts

Both study documents presented information exclusively in text form (e.g. no pictures were included). The experimental text (Appendix 1) was developed collaboratively by several highly experienced family physicians and general internists with expertise in colorectal cancer screening, including three of the current study authors (A.J., R.L.K., P.F.). The text was developed with the goal of presenting information regarding CRC screening test options, benefits, potential harms, and practical inconveniences

(e.g. barriers) in a readily understandable fashion. To facilitate comprehension, an effort was made to write the text at the lowest Flesch–Kincaid reading grade level possible, but without loss of critical meaning or context. For example, if the developers felt that removing or substituting a simpler alternative to a polysyllabic word (e.g. colonoscopy) that had been clearly and simply defined earlier in the text would hinder rather enhance patient comprehension of a subsequent passage, the more difficult word was retained, despite the resulting higher reading grade level in that passage. To further aid comprehension, most information in the experimental text was presented in a series of three tables (available test options, screening risks, and screening inconveniences), each with separate columns for FOBT and colonoscopy. Only three sections (the introduction, the passage on potential screening benefits, and the brief conclusion) were presented in non-tabular format.

We employed the Flesch–Kincaid method as the primary method of reading grade level estimation since it is available within Microsoft Word (Microsoft Corporation, Redmond, WA), facilitating the iterative development of the experimental text [31]. The method yields an estimated reading grade level based on the number of words per sentence (sl) and the number of syllables per word (spw), per the following formula embedded in Microsoft Word:

$$\text{Flesch–Kincaid reading grade level} = 0.39\,\text{sl} + 11.8\,\text{spw} - 15.59$$

As recommended by literacy experts, to obtain a more reliable estimate of text reading grade level, we applied the method to three different unique (i.e. non-overlapping) text samples, selected at random from within each text. The three reading grade level estimates were averaged to yield the summary estimate for each text [4]. Using this method, the estimated Flesch–Kincaid reading grade level for the experimental text was 7.4 (range of estimates 4–12).

The control text (Appendix 2), which had a Flesch–Kincaid reading level of 9.6 (range of estimates 9.5–9.8), was a hard copy version of the CRC screening "Fact Sheet," freely available on the National Cancer Institute web site at the time the study was conducted. The original text was modified for use in the RCT only by removing superfluous web navigational instructions. Beyond having a higher reading grade level than the experimental text, the control text presented less information in table format, employing only one brief table of FOBT and colonoscopy advantages and disadvantages. The control text also included less information on practical patient barriers to screening, and more information regarding CRC epidemiology and individual susceptibility risks.

Reading grade level estimates are known to vary among different assessment methods, with no consensus regarding the optimal method [4]. To help ensure that the experimental text reading grade level was actually lower than for the control text, we determined text reading grade levels via two additional methods, each applied to the same three randomly selected samples of each text employed in Flesch–Kincaid calculations, to yield average estimates. The first was the SMOG method, which is based on the total number of polysyllabic words contained in 30 sentences from a text. The SMOG reading grade level is determined by taking the square root of the nearest perfect square to the total number of polysyllabic words and then adding three to the resulting number [32]. The SMOG reading grade level estimate for the experimental text was 9.4 (range of estimates 7.6–12.5), and for the control text 10.7 (range of estimates 9.1–11.9). The second additional reading grade level method employed was the Fry formula, which measures sentences per 100 words and syllables per 100 words. For each text, the data from the three randomly sampled text passages was averaged, and plotted on a "readability graph" (y axis = average number of sentences per 100 words, x axis = average number of syllables per 100 words) to obtain a single estimate [33]. The Fry reading grade level estimate was 8.5 for the experimental text and 13 for the control text.

2.3. Measures

Recognition memory was measured using a signal detection theory-grounded measure, developed using a standard approach [27–29]. Briefly, the opening and closing sentences in both study texts were first eliminated, as were sentences with a word length more than two standard deviations from the document mean. Next, 20 undergraduate psychology students, recruited from the UCD Psychology subject pool and given course credit for participation, read both documents, highlighting in each the 20 sentences they felt were most critical to understanding the major points. The 20 most frequently highlighted sentences from each text comprised the final recognition memory test (40 sentences total).

In the RCT, patients were asked to indicate whether each sentence in the test was old or new. All of the recognition-memory test items (sentences) – including those which had appeared in the subject's randomly assigned text (old items) and those which had not appeared in the text (new items) – were phrased as true statements. This approach was employed to avoid respondent identification of false statements based solely on educated guesswork and/or pre-existing knowledge of CRC screening, rather than recognition that the statements had not appeared in their assigned text. Correctly identified old statements were recorded as hits [H], a measure of sensitivity, while new statements incorrectly identified as old were recorded as false alarms [F], a measure of response bias. A summary recognition-memory d prime (d') score was calculated using the formula

$$d' = z[H] - z[F]$$

where z is the transformation into a standardized z-score function that takes into account the number of possible hits or false alarms. This scoring method was appropriate to our analyses, which sought to isolate the effects of the studied texts themselves on recognition memory ("signal"), minimizing confounding by background topical knowledge and educated guesswork ("noise"). The range of d' scores in our study sample was −0.91 to 5.34.

Self-assessed health literacy was measured with a single item: "How often do you need to have someone help you when you read instructions, pamphlets, or the written material from your doctor or pharmacy?" Response choices were: 1 – never, 2 – rarely, 3 – sometimes, 4 – often, or 5 – always, with higher scores indicating more perceived health literacy problems. Prior cross-sectional studies suggested the item had reasonably good negative and especially positive predictive values [17–19]. *Socio-demographic characteristics* measured were age; ethnicity (Hispanic or non-Hispanic); race (White, Black, or Other); and education level (high school graduate, some college, college graduate [e.g. bachelor's degree], any post-graduate education [e.g. master's or doctoral coursework or degree]).

2.4. Analyses

Data were analyzed using PASW Statistics 18 (IBM, Armonk, NY). Independent-samples t-tests and chi-square tests were performed to assess the differences in baseline characteristics between experimental and control group patients. Linear regression was employed to examine the association of randomly assigned study text (the key independent variable) with recognition memory (the dependent variable), adjusting for age, education level, and self-assessed health literacy level. To further explore

education and health literacy influences on recognition memory, two subgroup analyses were also conducted. The first was limited to participants with high education (college degree or greater). The second was limited to participants with no self-reported health literacy problems (i.e. those responding "never" to the single item measure). We report standardized beta (β) regression coefficients.

3. Results

Fig. 1 shows the flow of participants through the study. One experimental group subject with missing education data was excluded from the analyses, leaving an analytic sample of 59 patients (29 experimental, 30 control). Table 1 shows the baseline characteristics of the analytic sample by study group. Participants were predominantly non-Hispanic white and well-educated, with high self-assessed health literacy. The correlation between education level and self-assessed health literacy was not statistically significant (Spearman $r = -0.22$, 95% confidence interval [CI]: −0.45, 0.04, $P = 0.10$). Only education level differed significantly between groups.

The mean unadjusted d' score was higher in the experimental group (2.54 versus 1.09, $t = -3.63$, $P = 0.001$). Table 2 shows the results of a linear regression of data from the analytic sample comparing d' score in experimental and control participants, adjusted for age, education level, and health literacy. The d' score was significantly higher in the experimental group. There was no significant relationship between the d' score and either education level or self-assessed health literacy. The experimental group also had a significantly greater d' score in analyses limited to those with a college degree or greater education ($\beta = 0.52$, 95% CI: 0.18, 0.86, $P = 0.004$), and to those with no self-reported health literacy problems ($\beta = 0.39$, 95% CI: 0.07, 0.71, $P = 0.02$).

4. Discussion and conclusion

4.1. Discussion

Prior RCTs examining the effects of different health information texts on patient literal comprehension did not concern the topic of CRC screening, and did not explore the potential influences of education or self-assessed health literacy [5–11]. The prior RCTs also did not employ signal detection theory-grounded recognition memory measures, to help disentangle the effects of study texts on literal comprehension from the effects of background topical knowledge, educated guessing, and systematic response bias. In the current study, confirming our first hypothesis, an experimental CRC screening text, written at a lower reading grade level, focused more on practical screening barriers, and presenting more information in comparative table format than a control text, fostered better recognition memory, assessed using a robust signal detection theory-grounded measure.

One factor limiting interpretation of the mechanism of this finding is that the study texts differed along several dimensions that might influence comprehension. Most notably, the experimental text had a lower reading grade level; was focused more on addressing practical barriers to screening (rather than CRC epidemiology and personal susceptibility as in the control text); and presented more information in tables comparing aspects of FOBT and colonoscopy (see Appendices 1 and 2 for details). However, our two study documents were more alike than the documents compared in most prior RCTs in this realm [5,6,8,10,11]. For example, in one of the RCTs, almost half of experimental document consisted of instructional drawings with minimal accompanying text, while the control document was all text (no illustrations) [5]. We also did not assess how much of the

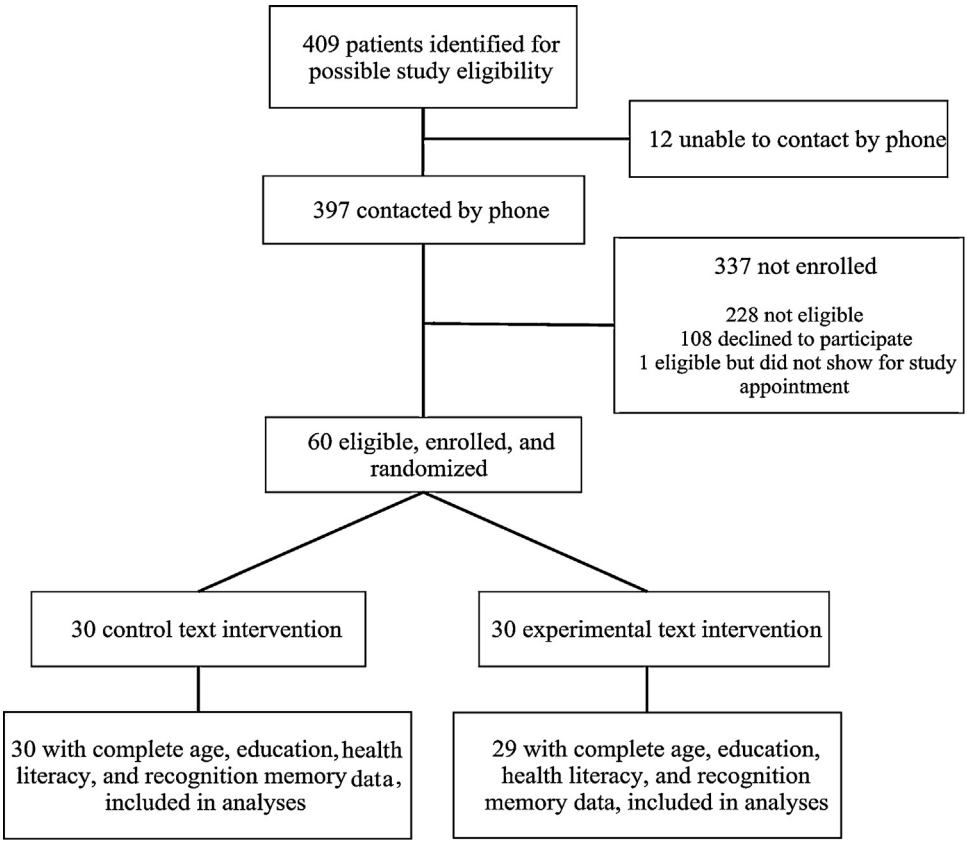


Fig. 1. Flow of participants through the study.

Table 1
Characteristics of the study sample.

Characteristic	Experimental group ($N=29$)	Control group ($N=30$)	P value
Age, years, mean (SD)	59.2 (6.0)	62.0 (7.1)	0.10[a]
Non-Hispanic White ethnicity, no. (%)[c]	21 (72.4)	19 (63.3)	0.57[b]
Race[c]			0.30[b]
White	22 (75.9)	21 (72.4)	
Black	1 (3.4)	5 (17.2)	
Other	6 (20.7)	3 (10.4)	
Highest education level, no. (%)			0.004[b]
High school graduate	–	3 (10)	
Some college	11 (37.9)	8 (26.7)	
College graduate	3 (10.3)	13 (43.3)	
Any post-graduate education	15 (51.7)	6 (20.0)	
Self-assed health literacy level,[d] no. (%)			0.50[b]
1 (highest)	18 (62.1)	20 (66.7)	
2	9 (31.0)	7 (23.3)	
3	1 (3.4)	3 (10.0)	
4	1 (3.4)	–	
5 (lowest)	–	–	

[a] Independent samples t-test.
[b] Chi-square test.
[c] Control group $N=29$ for ethnicity and race, due to missing data.
[d] Self-assessed with a single item, asking how frequently help is needed with reading instructions, pamphlets, or the written material from their doctor or pharmacy; 1 = never, 2 = rarely, 3 = sometimes, 4 = often, 5 = always.

study texts were actually read by participants. Thus, lower recognition memory among controls could have resulted from incomplete reading of the text, poorer recognition of sentences that were read, or both. Future studies that systematically manipulate and compare the effects of a range of document properties on patient comprehension, and assess which passages are actually read, would provide more precise guidance for designing comprehensible health information documents.

Contrary to our second hypothesis, we found evidence of experimental text benefits even in those with at least a college education and those with no self-assessed health literacy problems. One explanation is that texts with the features of our experimental text foster better literal comprehension across education and health literacy levels. If this were so, it would suggest that texts comparable to our experimental text may also facilitate inferential (higher level) comprehension across education and health literacy levels, since recognition memory is predictive of inferential comprehension [21–24]. If greater comprehension of such texts were shown to have a favorable influence on CRC screening decisions and behavior (e.g. increased uptake of screening), then providing texts like ours to all patients eligible for screening might be useful, regardless of patient education and self-assessed literacy levels. Such an approach, analogous to "universal precautions" in infectious disease prevention, already has been proposed by some experts [34].

Table 2
Results of a linear regression analysis of experimental text effects and participant age, education, and self-assessed health literacy effects on recognition memory d prime score.

Model variable	Estimated standardized β coefficient (95% CI)	P value
Experimental text (versus control)	0.42 (0.17, 0.68)	0.001
Age	−0.12 (−0.37, 0.14)	0.35
Education	−0.07 (−0.33, 0.19)	0.59
Self-assessed health literacy	−0.06 (−0.32, 0.19)	0.61

However, our study was limited in having few participants with low education and self-reported health literacy. RCTs involving patient samples representing a broader range of education and health literacy levels are needed to more rigorously examine the utility of a "universal precautions" approach.

We employed a single item self-report health literacy measure, previously validated only in cross-sectional analyses [17–19]. The measure was chosen to explore the potential utility of simple measures in time-constrained clinical settings, as recommended in quality improvement blueprints [1]. The lack of a statistically significant association between education level and self-reported health literacy in our study could have arisen if some study patients over-estimated their actual health literacy, questioning the utility self-reports to screen for health literacy problems in general practice. Different findings in regard to the influence of health literacy on comprehension might be observed in studies employing different health literacy measures, such as those based reading ability (e.g. Rapid Estimate of Adult Literacy in Medicine [REALM], Short Test of Functional Literacy in Adults [S-TOFHLA]) [4]. However, we believe that our choice of health literacy measure may not have been a critical issue driving our findings, given significant recognition memory benefits of the experimental text even in the analysis limited to persons with a college degree. This group, representing the top 30% of the educational hierarchy in the U.S. [35], would be anticipated to include the majority of "high health literacy" persons.

Nonetheless, additional larger RCTs are needed to more definitively examine the effects of different health information texts, not only on literal comprehension but also on *health behavior*, and moderation of these effects by health literacy and education. Ideally such RCTs should involve large and socio-demographically diverse patient samples, with a range of education and health literacy levels represented; address an array of different health topics and linked behavioral outcomes, including but not limited to CRC screening; and employ a variety of health literacy measures [4,36]. Of note, the unclear generalizability of our study findings to different patient subgroups, health topics, and health literacy measures, and their uncertain implications for health behavior, are limitations shared with previous RCTs in this realm [5–11].

We also noted considerable differences in the experimental and control text reading grade levels yielded by three widely employed methods (Flesch–Kincaid, SMOG, and Fry) [4]. Consistent with prior work, Flesch–Kincaid reading grade levels were lower than those obtained via the SMOG and Fry methods [37]. Also echoing prior research, *within* each of the methods, reading grade level estimates varied widely among different passages (samples) from a given text [4]. These findings underscore that estimated reading grade level, regardless of assessment method, is ultimately a crude measure of text complexity, and subject to limited reliability. Further research is needed to determine which specific properties of texts should be modified to improve literal comprehension. A number of text properties have been associated with literal comprehension in psycholinguistic studies conducted outside the healthcare realm with healthy volunteers, including word frequency (i.e. commonness of use in daily life), syntactic complexity, and semantic overlap (the degree to which subject content is related among sentences) [38–40]. RCTs involving *patient* samples, comparing literal comprehension and health behavior effects of documents prepared to differ systematically along such specific text properties, could provide further guidance for improving health information texts.

4.2. Conclusion

Compared with patients exposed to a control CRC screening information text, those exposed to an experimental CRC screening

text that was written at a lower Flesch–Kincaid reading grade level, focused more on addressing practical screening barriers, and employed more comparative tables had higher text recognition memory scores. Further, the salutary effect of experimental text exposure on recognition memory also was present in analyses limited to patients with least a college education and those with no self-assessed health literacy problems.

4.3. Practice implications

Providing CRC screening information texts with the features incorporated into our experimental text may be warranted for all screening-eligible patients, including those with relatively high education and self-assessed literacy levels, if such texts are eventually shown to improve CRC screening uptake.

Conflict of interest

The authors have no conflicts of interest to declare.

Role of funding

This work was funded in part by a National Cancer Institute Grant (R01CA131386, A.J.); a California Academy of Family Physicians Research Externship Grant (T.R.); and a UC Davis Department of Family and Community Medicine Research Grant (A.J.). The funders had no role in the study design; the collection, analysis and interpretation of data; the writing of the report; or the decision to submit the paper for publication.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.pec.2013.03.008.

References

[1] Committee on Health Literacy, Board on Neuroscience and Behavioral Health. Health literacy: a prescription to end confusion. Washington, DC: National Academies Press; 2004.
[2] Ad Hoc Committee on Health Literacy for the Council of Scientific Affairs, American Medical Association. Health literacy: report of the council on scientific affairs. J Amer Med Assoc 1999;281:552–7.
[3] Dewalt DA, Berkman ND, Sheridan S, Lohr KN, Pignone MP. Literacy and health outcomes: a systematic review of the literature. J Gen Intern Med 2004;19: 1228–39.
[4] Friedman DB, Hoffman-Goetz L. A systematic review of readability and comprehension instruments used for print and web-based cancer information. Health Educ Behav 2006;33:352–73.
[5] Davis TC, Bocchini Jr JA, Fredrickson D, Arnold C, Mayeaux EJ, Murphy PW, et al. Parent comprehension of polio vaccine information pamphlets. Pediatrics 1996;97:804–10.
[6] Davis TC, Fredrickson DD, Arnold C, Murphy PW, Herbst M, Bocchini JA. A polio immunization pamphlet with increased appeal and simplified language does not improve comprehension to an acceptable level. Patient Educ Couns 1998;33:25–37.
[7] Eaton ML, Holloway RL. Patient comprehension of written drug information. Am J Hosp Pharm 1980;37:240–3.
[8] Davis TC, Holcombe RF, Berkel HJ, Pramanik S, Divers SG. Informed consent for clinical trials: a comparative study of standard versus simplified forms. J Natl Cancer Inst 1998;90:668–74.
[9] Meade CD, Byrd JC, Lee M. Improving patient comprehension of literature on smoking. Am J Public Health 1989;79:1411–2.
[10] Coyne CA, Xu R, Raich P, Plomer K, Dignan M, Wenzel LB, et al. Randomized, controlled trial of an easy-to-read informed consent statement for clinical trial participation: a study of the Eastern Cooperative Oncology Group. J Clin Oncol 2003;21:836–42.
[11] Murphy DA, Hoffman D, Seage 3rd GR, Belzer M, Xu J, Durako SJ, et al. Improving comprehension for HIV vaccine trial information among adolescents at risk of HIV. AIDS Care 2007;19:42–51.
[12] US Preventive Services Task Force. Screening for colorectal cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2008;149:627–37.
[13] Davis TC, Dolan NC, Ferreira MR, Tomori C, Green KW, Sipler AM, et al. The role of inadequate health literacy skills in colorectal cancer screening. Cancer Invest 2001;19:193–200.
[14] Dolan NC, Ferreira MR, Davis TC, Fitzgibbon ML, Rademaker A, Liu D, et al. Colorectal cancer screening knowledge, attitudes, and beliefs among veterans: does literacy make a difference? J Clin Oncol 2004;22:2617–22.
[15] Peterson NB, Dwyer KA, Mulvaney SA, Dietrich MS, Rothman RL. The influence of health literacy on colorectal cancer screening knowledge, beliefs and behavior. J Natl Med Assoc 2007;99:1105–12.
[16] Jerant AF, Fenton JJ, Franks P. Determinants of racial/ethnic colorectal cancer screening disparities. Arch Intern Med 2008;168:1317–24.
[17] Morris NS, MacLean CD, Chew LD, Littenberg B. The Single Item Literacy Screener: evaluation of a brief instrument to identify limited reading ability. BMC Fam Pract 2006;7:21.
[18] Chew LD, Griffin JM, Partin MR, Noorbaloochi S, Grill JP, Snyder A, et al. Validation of screening questions for limited health literacy in a large VA outpatient population. J Gen Intern Med 2008;23:561–6.
[19] Jeppesen KM, Coyle JD, Miser WF. Screening questions to predict limited health literacy: a cross-sectional study of patients with diabetes mellitus. Ann Fam Med 2009;7:24–31.
[20] Fletcher JM. Measuring reading comprehension. Sci Stud Read 2006;10:323–30.
[21] Sheng HJ. A cognitive model for teaching reading comprehension. English Teach Forum 2000;38:12–5.
[22] Royer JM, Cunningham DJ. On the theory and measurement of reading comprehension. Technical report no. 91. Center for the Study of Reading, University of Illinois at Urbana-Champaign; 1978. http://www.ideals.illinois.edu/bitstream/handle/2142/17663/ctrstreadtechrepv01978i00091_opt.pdf?sequence=1 [accessed 20.12.12].
[23] Van den Broek P, Gustafson M. Comprehension and memory for texts: three generations of reading research. In: Goldman SR, Graesser AC, Van den Broek P, editors. Narrative comprehension, causality, and coherence. Mahwah, NJ: Erlbaum; 1999.
[24] Pettit NT, Cockriel IW. A factor study of the literal reading comprehension test and the inferential reading comprehension test. J Lit Res 1974;6:63–75.
[25] Paulhus DL. Measurement and control of response bias. In: Robinson JP, Shaver PR, Wrightsman LS, editors. Measures of personality and social psychological attitudes. San Diego, CA: Academic Press; 1991.
[26] Abdi H. Signal detection theory (SDT). In: Salkind N, editor. Encyclopedia of measurement and statistics. Thousand Oaks, CA: Sage; 2007.
[27] Long DL, Prat CS. Memory for Star Trek: the role of prior knowledge in recognition revisited. J Exp Psychol Learn Mem Cogn 2002;28:1073–82.
[28] Long DL, Wilson J, Hurley R, Prat CS. Assessing text representations with recognition: the interaction of domain knowledge and text coherence. J Exp Psychol Learn Mem Cogn 2006;32:816–27.
[29] Keating P. D-prime (signal detection) analysis. Los Angeles, CA: University of California Los Angeles, Phonetics Lab; 2004. http://www.linguistics.ucla.edu/faciliti/facilities/statistics/dprime.htm [accessed 09.02.12].
[30] Schulz KF, Grimes DA. Generation of allocation sequences in randomised trials: chance, not choice. Lancet 2002;359:515–9.
[31] Kincaid JP, Fishburne RP, Rogers RL, Chissom BS. Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy-enlisted personnel. Research Branch Report 8-75. Memphis, TN: U.S. Naval Air Station; 1975.
[32] McLaughlin GH. SMOG grading – a new readability formula. J Read 1969;12: 639–46.
[33] Fry E. A readability formula that saves time. J Read 1968;11:513–6.
[34] Paasche-Orlow MK, Schillinger D, Greene SM, Wagner EH. How health care systems can begin to address the challenge of limited literacy. J Gen Intern Med 2006;21:884–7.
[35] Bachelor's degree attainment tops 30 percent for the first time, Census Bureau reports. Washington, DC: U.S. Census Bureau; 2012, http://www.census.gov/newsroom/releases/archives/education/cb12-33.html [accessed 08.10.12].
[36] Baker DW. The meaning and the measure of health literacy. J Gen Intern Med 2006;21:878–83.
[37] D'Alessandro DM, Kingsley P, Johnson-West J. The readability of pediatric patient education materials on the World Wide Web. Arch Pediatr Adolesc Med 2001;155:807–12.
[38] Beck IL, McKeown MG, Sinatra GM, Loxterman JA. Revising social studies text from a text-processing perspective: evidence of improved comprehensibility. Read Res Q 1991;26:251–76.
[39] Haberlandt KF, Graesser AC. Component processes in text comprehension and some of their interactions. J Exp Psychol Gen 1985;114:357–74.
[40] Traxler MJ, Morris RK, Seely RE. Processing subject and object relative clauses: evidence from eye movements. J Mem Lang 2002;47:69–90.