# IMPROVING PATIENT COMPREHENSION OF NUMERIC HEALTH INFORMATION

Renato F. L. Azevedo, Daniel Morrow, Mark Hasegawa-Johnson, Kuangxiao Gu, Dan Soberal, Thomas Huang, William Schuh , Rocio Garcia-Retamero

Our goal is to improve understanding and use of numeric information (e.g., clinical test results) provided through portals to Electronic Health Record (EHR) systems by older adults with diverse numeracy and literacy abilities. We help older adults understand this information by emulating in portal environments best practices from face-to-face communication. To do this, we are developing a computer-based agent (CA) that will use nonverbal cues (e.g. voice intonation, facial expressions) as well as words to convey affective and cognitive meaning of the numbers and improve patient comprehension of the clinical information. The present paper describes a pilot study designed to evaluate the appropriateness and effectiveness of audio-video messages of a physician delivering clinical test results. These messages will serve as a template for the development of the CA. Older adult pilot participants generally understood the gist of the test results presented in the video messages. Participants' affective responses to the messages were appropriate to the message's level of risk: as the level of risk associated with the test results increased, positive affect decreased and negative affect increased. In addition, participants also thought the physician's delivery in the message matched the message content, and they thought that the messages were informative. These findings will be leveraged to finalize the materials for the primary study in which the impact of video and CA-based messages on patient comprehension of numeric information will be evaluated relative to standard formats used in patient portals.

## INTRODUCTION

Learning is vital to health across the lifespan, in part because patients are increasingly responsible for self-care. This is especially the case for older adults, who have more chronic illness, yet fewer cognitive resources for self-care (Morrow & Czaja, 2015). Because learning is often mediated by technology, success of learning for self-care depends on whether these environments are designed for older adults. Health information technology (HIT) has the potential to improve patient learning by enabling patients to access health information outside the constraints of brief clinical encounters. In particular, information provided through patient portals to Electronic Health Record (EHR) systems are intended to bridge patient self-care at home and primary care providers by ensuring ready access to health information and services. While this access may eventually revolutionize patient care, portals now function more as information repositories than as tools for patient education, and are often underutilized in part because they eliminate the clinical context that can support comprehension. More generally, they are not always designed with the needs and goals of older adults in mind. (IOM, 2012; Morrow & Chin, 2012; Taha et al., 2012).

Limited numeracy among patients is an important barrier to portal use. For example, federal 'meaningful use' EHR guidelines require health organizations to provide patients results from lipid panel, blood glucose, MRI, and other complex tests through portals (IOM, 2012). Although these requirements mandate providing patients with this information, there is no corresponding guidance about effective formats for communicating the information to diverse patients. Not surprisingly then, limited numeracy is an important predictor of portal use among older adults, even when controlling for cognitive ability (Morrow & Chin, 2012; Taha, et al., 2012). In short, technology enables ready access to health information but does not guarantee that the information is cognitively accessible. The goal of our project is to improve comprehension of numeric health information provided through portals.

## FUZZY-TRACE THEORY: SUPPORTING GIST-BASED UNDERSTANDING OF NUMERIC TEST RESULTS

Effective presentation of portal information should be guided by theory-based analysis of how older adults understand and reason about quantitative information. Fuzzy-trace theory suggests that verbatim and gist representations are encoded in parallel, stored separately and retrieved independently (Rivers, Reyna and Mills, 2008). The theory postulates that the simplest representations of choice ("fuzzy" memory traces) are more reliably stored in memory and preferred in retrieval, compared to more detailed and precise quantitative representations (verbatim memory traces) (Reyna et al., 2009). Gist representations are organized around affective/evaluative cognitive dimensions, and so may be important for motivating behavior as well as for comprehension. This analysis suggests that patients must do more than remember specific numbers to understand their test results. They integrate cognitive meaning ('How great is my risk of stroke given my family history?') and affective meaning ('Is my health threatened?') into a gist that captures the 'bottom-line' for their health and what to do about it. Gist-based understanding may be especially important for older adults, who rely on experience and affect to understand and reason about health information, perhaps to offset age-related limits in numeracy as well as sensory and cognitive declines (Peters et al., 2009; Reyna, 2011). However, while gist representations may better support decision making compared to verbatim representations, they can be challenging to create (Goldman et al. 2006). For example, the comprehension of lipid panel results requires understanding several component results with different scales. A higher "low density lipoproteins" (LDL) score means more risk, but a higher "high density lipoproteins" (HDL) score means less risk, and it requires integrating the components information into a global interpretation (high HDL protects against higher LDL). This may help explain why adults generally know little about their current or target cholesterol levels (Cheng et al. 2005; Nash et al., 2007).

Clinicians are traditionally a key to helping patients create gist representations by providing context for interpreting numeric information. They do so through verbal and nonverbal strategies that help older adults act on as well as understand information, which can improve health behaviors and outcomes (Ambady et al., 2002). Unfortunately, clinicians do not consistently use these strategies because of limited communication time and training. Currently, portals may exacerbate this problem of limited provider support for patient comprehension because they present numeric information to patients with little context.

## IMPROVING PATIENT COMPREHENSION OF NUMERIC INFORMATION FROM PORTALS

The goal of our project is to improve the use of portal information by older adults with diverse numeracy and literacy abilities, so that portals can better support patient-centered care. We will take advantage of computer-based agents (CA) in order to emulate in portal environments best practices in face-to-face communication that help patients understand health information at a gist level. Members of our team have developed a CA that uses nonverbal cues such as prosody (stress and voice intonation), facial expressions, as well other nonverbal cues to convey affective as well as cognitive meaning (Huang et al., 2009).We will refine this CA so that it serves as an expert clinical portal intermediary that provides succinct commentary on risk information. The verbal and nonverbal cues in this multimedia format reinforce each other, and may offset the impact of age-related sensory and cognitive declines on comprehension and decision making (Van Gerven, Pass et al., 2003). CA nonverbal cuing of affective information may especially support older adults' intention to act because elders often rely on affective information to make health decisions (Peters et al., 2009). CAs are more likely than text interfaces to engender social responses such as trust to automated systems. While CAs are used in many education settings, there have been few applications to health care (but see Bickmore et al., 2010).

In order to create the CA, we first video-recorded a physician in order to provide a template for CA development. These audio-videos will help refine the existing CA. The first planned experiment will compare comprehension and reactions to clinical test result messages presented in the standard, verbally enhanced, graphically enhanced, , with the video-recorded physician serving as proxy for the CA. Later experiments will compare video and CA-enhanced formats. We will compare the impact of standard portal formats (text and numeric) to several enhanced formats on comprehension of lipid and blood glucose test results. In the verbally enhanced format, labels for evaluative categories (more or less risk) that provide context for interpreting the specific numbers will be included. These labels also promote emotional processing of the quantitative information (Peters et al., 2009). They may be especially helpful to older adults, who tend to have high levels of verbal ability (Schwartz, et al. 1997) and focus on emotional meaning (Charles & Carstensen, 2010). In the graphically enhanced condition, graphics that convey key relational features (larger/smaller than) that support gist understanding and of risk will be included (Garcia-Retamero & Galesic, 2009). Finally, in the CA condition, the same graphics will be accompanied by a CA that provides commentary about the information, with nonverbal cues (prosody, facial expressions) signaling information relevance and guiding affective interpretation, as in ideal face-to-face communication. The graphics and numbers are also included in this condition, with visual cues that link CA commentary with relevant information in the graphic. The multimedia CA format should be most effective because the verbal and nonverbal cues reinforce each other (Van Gerven, Pass et al., 2003).

Before conducting the primarystudy, we conducted a pilot study to ensure that the video recorded messages were appropriate for developing a CA for older adults. We evaluated four questions related to the appropriateness and effectiveness of the audio-video messages of a physician delivering clinical test results:

1. Do older adults understand the gist of the messages?
2. Are participants' affective responses appropriate to the message content (lev-

el of risk associated with the test results)?
3. Do participants think the physician's delivery is appropriate for the clinical information that is conveyed?
4. Do the participants think the messages are informative?

## METHOD

**Participants.** Participants were 12 community-dwelling older adults (average age of 76.6 years, range 65-89 years; 66.67% females), native English speakers, with no physical, cognitive or visual/auditory impairments that could limit participation. They provided consent before participation.

**Materials.** Twenty-four messages that conveyed results from cholesterol and from HbA1c diabetes screening test results were developed in collaboration with two physicians of our partner health care system to ensure realistic clinical content. Messages for diabetes screening test results as well as cholesterol results were developed because type 2 diabetes is a common age-related chronic illness that is often comorbid with cholesterol-related illness. The cholesterol messages described complex patterns of scores on multiple tests (total cholesterol, triglycerides, high density lipoproteins (HDL), and low density lipoproteins (LDL)) that suggested low, borderline, or high risk for cardio-vascular illness. To help patients understand the overall risk associated with each message, the message ended with a summary of the overall risk for heart disease associated with the test scores. There were equal numbers of cholesterol and diabetes messages reporting test results from each level of risk. Because these risk levels depend on patient-specific risk factors (e.g., coronary artery disease, hypertension, family history of heart disease) as well as the scores, we also created patient profiles for each message. The 12 cholesterol and 12 diabetes messages were presented in blocks. Within each block there were four messages at each three risk levels (low, borderline, and high).

The messages were recorded by a retired physician, who presented the information as if he were discussing the results and their implications with his patient. Each recorded message was divided in two segments: (a) description of all test result components and the associated level of risk for each component, and (b) an overall summary of risk. In the script for the messages, important information was italicized and the most important/relevant information was bolded to indicate what the physician should emphasize when talking to the patient.

| |
|---|
| "I'm going to tell you about the results of your cholesterol test. Your *total cholesterol score* is *145*. This is a ***good score*** *and shows that you have* **low** ***risk for heart disease***. *Your triglycerides* score is *148*. This score is ***normal, and is another sign that you have low risk for heart disease***. Your *HDL or good cholesterol is 43*. This score is ***borderline. A higher HDL score is desirable*** *because these lipoproteins help remove cholesterol from your bloodstream and artery walls, which lowers your risk.* ***In addition****, your LDL, or bad cholesterol, score is 72*. This score is also good in terms of your *risk for heart disease*. ***Overall*****, *I feel your risk for heart disease is low. Your test results are good so there is no need to worry about your results at this time.*** ." |

*Table 1.* Example of cholesterol message

**Measures.**

**Demographics.** Participants completed a questionnaire about their age, gender, ethnic background and education.

**Questions about messages**

**1. Risk Perception.**

Gist comprehension of risk associated with both individual test scores (e.g., HDL) and with the complete message was evaluated. We probed overall risk before as well as after the second part of the message (summary statement) in order to evaluate how well participants could extract overall gist only from the component scores. In addition to overall risk perception, the additional questions probed participants' reactions to each message (ordinal level gist: low/borderline/high; Reyna et al. 2009).

**2. Affective reactions.** Participants indicated their affective reaction to the messages. They indicated to what extent they experienced 7 negative and 7 positive emotions by responding, for each emotion, to a 9-point scale ranging from 1 (not at all) to 9 (very much), as follows: "*If you were the patient Sam, how would you feel as you watched this message? Indicate the extent that you felt:* (*assured, calm, cheerful, happy, hopeful, relaxed, and relieved*; or *anxious, afraid, discouraged, disturbed, sad, troubled, and worried" (Garcia-Retamero & Cokely, 2011).*

**3. Physician Delivery.** Participants indicated the extent to which they thought the physician delivery was appropriate for the clinical information that was conveyed (level of risk associated with test results), on a 9-point scales ranging from 1 (strongly disagree) to 9 (strongly agree), as follows: "*I think that the way the doctor conveyed the information matched what he said".*

**4. Satisfaction with the message.** Participants indicated the extent to which they considered the information conveyed in the messages was useful, on a 9-point scales ranging from 1 (not at all useful) to 9 (very useful), as follows: "*How useful do you think was the information conveyed in this message?" (Garcia-Retamero & Cokely, 2011).*

**Procedure.** Participants viewed all 24 video messages. Each message was divided into two parts: a) the physician first described the test results, followed by questions about the risk information in the message; b) Then, the summary statement was presented, followed by additional questions about the complete message's content and presentation.


*Figure 1.* Study procedure.

## PRELIMINARY RESULTS

Participants extracted the gist for an overall level of risk from the messages before hearing the summary risk statement (cholesterol: 82% correct, A1C: 83% correct), and hearing the summary statement improved this gist accuracy (cholesterol: 93 % correct, A1C: 85%). The average accuracy for each level of risk is presented in Table 2.

| | **Cholesterol scenarios** | | **Diabetes scenarios** | |
|---|---|---|---|---|
| | ***Accuracy – Before Summary Statement*** | ***Accuracy – After Summary Statement*** | ***Accuracy – Before Summary Statement*** | ***Accuracy – After Summary Statement*** |
| *Low risk* | 93.75% | 89.58% | 100.00% | 93.75% |
| *Borderline risk* | 70.83% | 91.67% | 50.00% | 60.42% |
| *High risk* | 81.25% | 97.92% | 97.92% | 100.00% |
| ***Total*** | **81.94%** | **93.06%** | **82.64%** | **84.72%** |

*Table 2.* Summary results for average accuracy gist comprehension in each associated level of risk.

The results also suggest that the summary statement improved understanding more for the more complex cholesterol messages than for the simpler diabetes messages. In addition, there was some evidence that messages with more complex test results (indicating borderline rather than consistently low or high risk) were more difficult to understand. As expected, there was some evidence that older adults had more trouble retrieving an accurate gist representation at the component level for more complex test results (e.g. cholesterol) than for the single score (e.g. A1C). For cholesterol test results, however older adults were able to retrieve correctly the gist representation of the total cholesterol score for most of the scenarios, 50% of the scenarios have low scores of accuracy for HDL, 67% of the scenarios have low scores of accuracy for LDL, and 50% of the scenarios have low scores of accuracy for Triglycerides.

Participants' affective responses to the messages were appropriate to the message's level of risk: as the level of risk associated with the test results increased, positive affect decreased and negative affect increased. Similar results were found for both the cholesterol and diabetes measures, as shown in Table 3.

| **Cholesterol scenarios** | **Diabetes scenarios** |
|---|---|
| ***Low risk*** | ***Low risk*** |
| Positive emotions: 8.48 | Positive emotions: 8.68 |
| Negative emotions: 1.51 | Negative emotions: 1.20 |
| ***Borderline risk*** | ***Borderline risk*** |
| Positive emotions: 4.98 | Positive emotions: 4.24 |
| Negative emotions: 4.63 | Negative emotions: 5.35 |
| ***High risk*** | ***High risk*** |
| Positive emotions: 2.70 | Positive emotions: 2.65 |
| Negative emotions: 6.22 | Negative emotions: 7.13 |

*Table 3.* Summary results for affective responses in each associated level of risk

Pilot participants also considered that the physician's message delivery (facial expressions and tone of voice) was appropriate for the content of the message (1-9 scale where 9 is very appropriate: mean of 8.6 for A1C messages; 8.5 for cholesterol messages). This result converges with the affective response findings to suggest that the nonverbal cues in the video helped convey the affective gist of the message. Participants also thought that the information conveyed by the messages was useful (1-9 scale: 8.6 for A1C, 8.4 for cholesterol).

## CONCLUSIONS

Older adults were generally able to understand the gist of the video test messages for both cholesterol and diabetes tests, especially after hearing the summary risk statement.

Equally important, participants' affective responses to the messages were appropriate to the message's level of risk: As the level of risk associated with the test results increased, positive affect decreased and negative affect increased. The same pattern occurred for the cholesterol and for the diabetes messages. Moreover, participants thought that the physician's delivery (e.g., tone of voice) was appropriate for the level of risk conveyed. These results suggest that participants developed a gist representation organized in terms of affective as well as cognitive dimensions. As discussed, because gist representations are organized around affective dimensions, they may be important for motivating behavior.

In the primary study we will explore in addition to questions about affective responses to the message, questions about behavioral intentions and attitudes toward self-care behaviors (e.g. use of medication, diet, exercise).

In the primary study weThese results suggest that the video-recorded physician is an appropriate template for develop the CA. They also encourage us to further explore formats that could enhance understanding of gist representations at the component level and most important the associations with an overall gist representation.

As a preliminary pilot study, the present findings are limited by a small sample size. However, our intent was not to generalize beyond the sample of participants (and especially not the one clinician), but rather to evaluate the appropriateness and effectiveness of the audio-video messages to serve as the basis for the CA.

## ACKNOWLEDGMENTS

Support for this research provided by AHRQ grant: R21HS022948.

## REFERENCES

Ambady, N., Koo, J., Rosenthal, R., & Winograd, C. (2002). Physical therapists' nonverbal communication predicts geriatric patients' health outcomes. Psychology and Aging, 17(3), 443-452.

Bickmore, T., Pfeifer, L., Byron, D., Forsythe, S., Henault, L., Jack, B., et al. (2010). Usability of conversational agents by patients with inadequate health literacy: Evidence from two clinical trials. Journal of Health Communication, 15, 197-210.

Charles, S. T., & Carstensen, L. L. (2010). Social and emotional aging. Annual review of psychology, 61, 383-409.

Cheng, S., Lichtman, J. H., Amatruda, J. M., Smith, G. L., Mattera, J. A., Roumanis, S. A., & Krumholz, H. M. (2005). Knowledge of cholesterol levels and targets in patients with coronary artery disease. Preventive cardiology, 8(1), 11-17.

Garcia-Retamero, R., & Cokely, E. T. (2011). Effective communication of risks to young adults: Using message framing and visual aids to increase condom use and STD screening. Journal of Experimental Psychology-Applied, 17(3), 270.

Garcia-Retamero, R., & Galesic, M. (2009). Communicating treatment risk reduction to people with low numeracy skills: A cross-cultural comparison. Journal Information, 99(12).

Goldman, R. E., Parker, D. R., Eaton, C. B., Borkan, J. M., Gramling, R., Cover, R. T., & Ahern, D. K. (2006). Patients' perceptions of cholesterol, cardiovascular disease risk, and risk communication strategies. The Annals of Family Medicine, 4(3), 205-212.

Huang, T. S., Hasegawa-Johnson, M. A., Chu, S. M., Zeng, Z., & Tang, H. (2009). Sensitive Talking Heads [Applications Corner]. Signal Processing Magazine, IEEE, 26(4), 67-72.

Institute of Medicine (2012). Health IT and Patient Safety: Building Safer Systems for Better Care. Washington DC: The National Academies Press.

Morrow, D.G. & Czaja, S.J. (2015). The implications of aging for Human-System Integration. In Deborah Boehm-Davis, Francis Durso, and John D. Lee (Eds), Handbook of Human-System Integration. Washington, DC: American Psychological Association. In press.

Morrow, D. G. & Chin, J. (2012). Technological bridges between health care systems and older adults. In R. Hill, R. Zheng, & M, Gardner (Eds.) Engaging older adults with modern technology: Internet use and information access needs (59-79). Hershey, PA: IGI Publishing.

Nash, I. S., Mosca, L., Blumenthal, R. S., Davidson, M. H., Smith Jr, S. C., & Pasternak, R. C. (2003). Contemporary awareness and understanding of cholesterol as a risk factor: results of an American Heart Association national survey. Arch Internal Medicine, 163(13), 1597.

Peters, E., Dieckmann, N.F., Västfjäll, D., Mertz, C.K., Slovic, P., & Hibbard, J. H. (2009). Bringing meaning to numbers: The impact of evaluative categories on decisions. Journal of Experimental Psychology: Applied, 15(3), 213–227.

Reyna, V. (2011). Across the life span. In B. Fischoff, N. Brewer, & J. Downs (Eds). Communicating Risks and Benefits: A Users Guide (pp 111-120). Washington DC: FDA.

Reyna, V. F., Nelson, W., Han, P., & Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. Psychological Bulletin, 135, 943-973.

Rivers, S. E.; Reyna, V. F.; Mills, B. (2008). Risk taking under the influence: a fuzzy-trace theory of emotion in adolescence. *Dev. Rev*, *28*, 107-144.

Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997). The role of numeracy in understanding the benefit of screening mammography. Annals of internal medicine, 127(11), 966-972.

Taha, J., Sharit, J., & Czaja, S. J. (2012). The Impact of Numeracy Ability and Technology Skills on Older Adults' Performance of Health Management Tasks Using a Patient Portal. Journal of Applied Gerontology.

Van Gerven, P. W.M., Paas, F., Van Merriënboer, J. J.G., Hendriks, M., & Schmidt, H. G. (2003). The efficiency of multimedia learning into old age. British Journal of Educational Psychology, 73, 489–505.