



Published in final edited form as:

Behav Res Methods. 2017 August ; 49(4): 1386–1398. doi:10.3758/s13428-016-0794-1.

Active Engagement in a Web-Based Tutorial to Prevent Obesity Grounded in Fuzzy-Trace Theory Predicts Higher Knowledge and Gist Comprehension

Priscila G. Brust-Renck¹, Valerie F. Reyna¹, Evan A. Wilhelms¹, Christopher R. Wolfe², Colin L. Widmer², Elizabeth M. Cedillos-Whynott², and A. Kate Morant¹

¹Cornell University

²Miami University

Abstract

We used Sharable Knowledge Objects (SKOs) to create an Intelligent Tutoring System (ITS) grounded in Fuzzy-Trace Theory to teach women about obesity prevention: GistFit, getting the gist of healthy eating and exercise. The theory predicts that reliance on gist mental representations (as opposed to verbatim) is more effective in reducing health risks and improving decision making (e.g., Reyna & Lloyd, 2006; Reyna & Mills, 2014). Technical information was translated into decision-relevant gist representations and gist principles (i.e., healthy values). The SKO was hypothesized to facilitate extracting these gist representations and principles by engaging women in dialogue, “understanding” their responses, and replying appropriately to prompt additional engagement. Participants were randomly assigned to either the obesity prevention tutorial (GistFit) or a control tutorial containing different content using the same technology. Participants were administered assessments of knowledge about nutrition and exercise, gist comprehension, gist principles, behavioral intentions and self-reported behavior. An analysis of engagement in tutorial dialogues and responses to multiple-choice questions to check understanding throughout the tutorial revealed significant correlations between these conversations and scores on subsequent knowledge tests and gist comprehension. Knowledge and comprehension measures correlated with healthier behavior and greater intentions to perform healthy behavior. Differences between GistFit and control tutorials were greater for participants who engaged more fully. Thus, results are consistent with the hypothesis that active engagement with a new gist-based ITS, rather than a passive memorization of verbatim details, was associated with an array of known psychosocial mediators of preventive health decisions, such as knowledge acquisition, and gist comprehension.

Correspondence concerning this article should be addressed to Valerie F. Reyna, Human Neuroscience Institute, G331 Martha Van Rensselaer Hall, Cornell University, Ithaca, NY, 14853, USA. vr53@cornell.edu.

Priscila G. Brust-Renck, Department of Human Development, Cornell University; Valerie F. Reyna, Departments of Human Development and Psychology, Human Neuroscience Institute, Cornell Magnetic Resonance Imaging Facility, and Center for Behavioral Economics and Decision Research, Cornell University; Evan A. Wilhelms, Department of Human Development, Cornell University; Christopher R. Wolfe, Department of Psychology, Miami University; Colin L. Widmer, Department of Psychology, Miami University; and Elizabeth M. Cedillos-Whynott, Department of Psychology, Miami University; A. Kate Morant, Department of Biological Sciences, Cornell University.

Keywords

Fuzzy-Trace Theory; Intelligent Tutoring System; Gist; Health; Obesity

We report the results of a new approach to cognitive engagement in preventing obesity, which is a serious and growing problem in the United States and indeed much of the world (Ogden, Carroll, Kit, & Flegal, 2014). The novel approach is a web-based tutoring system with a talking pedagogical agent that interacts with learners in their natural language developed based on fuzzy-trace theory (Reyna & Brainerd, 1995, 2011). We incorporate relevant research and theoretical mechanisms involved in active learning environments to create a comprehensive account of processes that guide decision-making. First, we present traditional ways to communicate information broadly through static websites and modern web-based tutors based on artificial intelligence. Then, we discuss the importance of theoretical underpinnings to efficient interventions, and how these lead to the current study of engagement in acquiring health knowledge, gist comprehension (beyond learning of literal facts), and endorsement of healthy values or principles.

The success of traditional communication approaches relies on the idea that pertinent information is communicated and encoded so that it can be retrieved later. In particular, these approaches suggest that communicating information requires presenting detailed and precise information while also targeting misperceptions to change people's beliefs and behaviors (Li & Chapman, 2013). An obesity intervention or prevention program might, for example provide instruction on how to read nutrition labels or how to calculate the amount of calories consumed each day. However, emphasizing the learning of verbatim facts is often insufficient to change either beliefs or behaviors (Brust-Renck, Royer, & Reyna, 2013; Peters, 2009; Reyna & Farley, 2006). A common result of such approaches is that behavior change is demonstrated in the short term, but outcomes are either not followed over the long term or effects dissipate over time (e.g., Cooper et al., 2010).

Few members of the public have sufficient knowledge to understand health and medical messages, such as those involving breast cancer risk, vaccination, HIV/AIDS, and obesity (e.g., Downs, Bruin de Bruine, & Fischhoff, 2008). Lack of background information to connect the dots (i.e., achieve global coherence) compromises comprehension and retention of messages (Lloyd & Reyna, 2009; Reyna & Adam, 2003). Those with the expertise to understand these messages and connect the dots, such as genetic counselors, are not widely available (as there is approximately one certified genetic counselor to every 100,000 people in the United States; ABGC, 2014) or are not always covered by insurance. Information available on the web and through social media does not necessarily solve this problem because the information is often not clear and easy to understand (in particular, from official government sites; Betsch et al., 2012; Reyna, 2012b).

One classic theoretical approach to bridge the gap between learned materials and reality comes from Gestalt psychology through urging reasoners to "Think!" in order to decrease mindless reactions to a task and promote active thinking about what would be the best approach (e.g., Reyna, Lloyd, & Brainerd, 2003). The mechanism behind this approach is the distinction between reproductive (or non-productive) and productive thinking. The

former is a result of rote associations (e.g., students who are shown how to measure the area of a rectangle tend to mistakenly apply the same formula to a parallelogram) and the latter is a result of conceptual understanding that supports transfer to new instances (e.g., by realizing that cutting off the triangle from one end of the parallelogram and placing it on the other side to form a rectangle, then the surfaces of both figures become similar and the area can be determined; Wertheimer, 1959; Wolfe & Reyna, 2010a, b; Wolfe, Fisher, Reyna, & Hu, 2012). According to Wertheimer, reproductive thinking is a recombination of learned associations (i.e., memorizing the content of the curriculum by forming associations among the words, much like Skinner described learning in rats), whereas productive thinking involves deep conceptual insight that supports meaningful inferences from learned materials to superficially different but conceptually similar examples. According to Gestalt psychology, information learned by productive thinking is better retained in memory (i.e., the essential features of learned material are remembered) than information learned by rote memorization (for a review, see Wertheimer, 2010).

Using this approach to encourage connections in learned information in such a way that influences behavior change has often required the same one-on-one tutoring, which is a gold standard for communicating information. Human tutors have the advantage of asking questions, they can immediately gauge progress and correct mistakes, and they can encourage students to elaborate on answers to questions of knowledge and comprehension. This elaboration has benefits to understanding. For example, elaboration on topics that people mistakenly think they understand has the effect of correcting confidence in understanding, particularly regarding concepts related to causal models (Fernbach, Rogers, Fox, & Sloman, 2013). More generally, actively generating explanations of material improves learning compared to passive reading or listening to lectures (Graesser, McNamara, & VanLehn, 2005). One promising direction has been that of automated tutoring systems to emulate the most effective strategies from one-on-one learning. As previously described, however, this can be expensive (i.e., it requires training) or otherwise unavailable (mostly due to location issues, time constraints, and scarcity of resources).

Progress has been made in creating automated instruction that can teach complex conceptual material with the efficacy of human tutors (Wolfe et al., 2013; Wolfe, Reyna, Widmer, Cedillos, et al., 2015). Intelligent Tutoring Systems (ITSs; i.e., computer-based tutoring) have been developed to facilitate human-computer interactions and overcome hurdles resulting from the lack of a dynamic teaching environment and trained teachers (Graesser et al., 2004). One of the greatest advantages of computer tutors is their ability to mimic human tutors in helping learners to connect the dots and provide immediate feedback, much like one-on-one tutors, facilitating the integration between current and prior experiences, which are often features of one-on-one tutoring (Chi, Siler, Jeon, Yamauchi, & Hausmann, 2005; Lloyd & Reyna, 2009).

ITSs can be programmed to elicit elaborate explanations from students and can communicate with students using natural language. Examples of platforms for systems that use semantic decomposition to interact with people in natural languages are AutoTutor and Sharable Knowledge Objects (SKOs, formerly Autotutor Lite; Hu, Han, & Cai, 2008; Hu, Cai, Han, Craig, Wang, & Graesser, 2009), which have been successfully applied to teaching

about computer science, physics, and genetic risk of breast cancer, among other topics (e.g., Graesser et al., 2004; Wolfe et al., 2013; Wolfe, Reyna, Widmer, Cedillos, et al., 2015). In these systems, a talking avatar communicates information using conversational language, facial expressions, graphical displays, and videos. This platform accomplishes a tailored interaction through the use of curriculum scripts including ideal answers, responses for common misconceptions, and feedback for the student (i.e., questions are answered in an efficient and effective manner). Moreover, people can engage in dialogues with the tutorial to actively generate explanations in more effective and deeper ways than when people are just given static information (Arnott, Hastings, & Allbritton, 2008; Chi, 2009).

This strategy on its own, however, does not necessarily employ findings from the latest theories regarding how to successfully change behavior as a result of learned information, though it may be a useful tool with which to implement theoretical predictions. BRCA Gist (BRCAst Gist Intelligent Semantic Tutoring), an ITS devised around fuzzy-trace theory, has been shown to lead to improvements in knowledge, comprehension, decision to undergo genetic testing based on one's personal risk, and assessment of risk from multiple scenarios in randomized, controlled experiments (Wolfe et al., 2013; Wolfe, Reyna, Widmer, Cedillos, et al., 2015). The web-based tutorial has multiple advantages as a method of instruction: it is asynchronous (can be accessed at any time of day and night), interactive, and multi-media (it transmits the information orally, in writing, and with figures and photographs). The tutorial allows users to interact with it through the use of natural language by asking deep-level reasoning questions and providing feedback, encouragement, and additional information to aid understanding. The tutorial attempts to comprehend participants' answers and to simulate replies from human tutors (VanLehn, 2011). In addition, results showed that not all active tutoring helped to improve outcomes, but the development of gist explanations (based on fuzzy-trace theory) was responsible for significant improvements (Wolfe, Reyna, Widmer, Cedillos-Whynott, et al., 2015).

Thus, we turn to fuzzy-trace theory, a theory influenced by the principles of Gestalt psychology, that emphasizes two processes of reasoning (i.e., gist and verbatim thinking) that reflect different types of information processing (Reyna, 2012a, 2013; Reyna & Brainerd, 1995, 2011; Reyna et al., 2003). Gist thinking involves bottom-line (meaningful) understanding of conceptual information, the "substance" that exists irrespective of exact words, numbers, or pictures (which supports productive thought). Verbatim (literal) thinking relies on the surface form of information, the exact mental representation of the stimulus (which does not support productive thought). Fuzzy-trace theory predicts independent and parallel processing of verbatim-based and gist-based thinking, as experiments have demonstrated (e.g., Brainerd, Reyna, & Howe, 2009; Reyna & Brainerd, 1992). In the context of communication, gist thinking relies on extracting the essential, bottom-line meaning of information, which results in insightful intuition, and retrieval of relevant knowledge and values, whereas verbatim thinking is a result of associative activation (i.e., mindless memorization of learned materials), which cannot predict far transfer of learning to novel instances.

In the present study, we investigated the role of active learning using a web-based ITS that applies artificial intelligence to create a scalable and cost-effective way to engage many

people in dialogue about obesity simultaneously. The novelty of our method for teaching involves combining fuzzy-trace theory with artificial intelligence. The ITS was developed to engage participants in a dialogue about the myriad issues associated with obesity prevention, from nutrition to exercise. Previous research has shown that generating arguments about a specific topic has been used to increase understanding (Chi, 2009; VanLehn, Graesser, Jackson, Jordan, Olney, & Rosé, 2007; Wolfe, Reyna, Widmer, Cedillos, et al., 2015). The prediction is that actively generating and elaborating on explanations of complex materials promotes understanding of bottom-line (gist) semantic meaning of information (Clariana & Koul, 2006; Lloyd & Reyna, 2001). Our methods take advantage of the fact that interactive dialogues with tutors providing immediate feedback increases meaningful processing of information that occurs during encoding and forges stronger memory traces of semantic meaning (Clariana & Koul, 2005, 2006; Reyna & Brainerd, 2011).

Moreover, our tutor was guided by theoretical principles from fuzzy-trace theory, particularly emphasizing active comprehension of bottom-line (gist) meaning from conceptual information and extraction of gist principles (i.e., healthy values). As mentioned previously, fuzzy-trace theory posits that meaningful understanding goes beyond the surface details and encourages connecting the dots between current and prior content (Reyna, 2013; Reyna et al., 2003). By supplying the “active ingredient” (i.e., the ability to extract gist) in an active environment (i.e., multiple-choice questions embedded within the tutorial and the engagement of participants in tutorial dialogues), we predicted that we would find an association between the level of engagement with our tutorial and knowledge and comprehension, as well as endorsement of healthy values (i.e., gist principles). We also hypothesized that these measures would be associated with intentions to perform healthier behavior in the future. That is, engagement should promote knowledge acquisition, gist comprehension, and endorsement of gist principles, which, in turn, have been shown to be psychosocial mediators of health behaviors and behavioral intentions (e.g., Reyna & Mills, 2014; Wolfe, Reyna, Widmer, Cedillos, et al., 2015).

Finally, we included measures of a traditional theory (i.e., Theory of Reasoned Action/Theory of Planned Behavior; Ajzen, 2011; Fishbein, 2008) to test alternative hypotheses about the possible benefits of the tutorial (i.e., beliefs, attitudes, and norms have been consistently linked to behavioral intentions and behavior). However, none of these variables was related to either group or tutorial engagement, and will therefore not be reported here.

Method

Design

We assessed the relationship between engagement and health-related psychosocial mediators by studying active learning through tutorial dialogues as people interacted with an obesity prevention ITS (GistFit). GistFit engagement was broken down into three levels: Participants engaged with the tutorial in either *none*, *some*, or *all* dialogues. The assessment of the dialogues was embedded in a larger randomized, controlled experiment of the effectiveness of GistFit in teaching women about obesity prevention. In contrast to the scope of this article, which focuses on the assessment of the tutorial interaction with participants, we also compared the efficacy of GistFit and a traditional tutorial about obesity prevention to

irrelevant-content control tutorials about genetic risk and breast cancer risk (Brust-Renck et al., 2012). These analyses showed that the GistFit intervention effectively increased behavioral intentions, knowledge, and gist comprehension compared to control groups. Participants were randomly assigned to either GistFit or one of the highly similar control groups about genetic risk and breast cancer risk; Gistfit and control groups did not differ on baseline characteristics by more than chance levels (Table 1).

Participants

Participants were 251 female undergraduate students recruited from the participant pools at two universities in the Midwest and Eastern United States; 25 did not complete the entire tutorial and 6 completed the survey quickly (i.e., in less than half of the average time, 8 from GistFit and 23 from control tutorials); analyses were conducted using an intent-to-treat approach not excluding these 31 participants. None of the “dropouts” were due to participants deciding to not complete the tutorial. Instead, they were due to a technical problem that affected all of the groups during the tutorial, and they were all included in the evaluation/assessments. All participants provided written consent, and the project was approved by each university’s Institutional Review Boards. Participants were recruited online (though the study was conducted in the laboratory in-person) and they either received course credit or volunteered for participation. The mean age of participants was 19.21 (SD = 1.73). Overall, 61.5% participants identified themselves as Caucasian, 10% as African American, 21.8% as Asian, and 6.7% as mixed/other race; 9.1% identified as Hispanic (see Table 1 for demographic characteristics for each tutorial). There were 37 participants in the GistFit condition and 183 were assigned to one of five highly similar control conditions with unrelated content (participants in each control condition ranged from 31 to 41). The control groups differed in minor ways, each using slightly different ways of presenting the same information in tutorials about breast cancer and genetic risk (i.e., two full versions with minor “back end” improvements, two versions without questions that require either explanation or argumentation, and one version without a few of the figures and video; Wolfe, Reyna, Widmer, Cedillos-Whynott, et al., 2015). As expected, the results across controls groups for the unrelated breast cancer curriculum were very similar. Thus, we treat the control conditions as a single group.

Materials

Tutorial—Our tutorial, GistFit, getting the gist of healthy eating and exercise (Brust-Renck et al., 2012), is an adapted version of EatFit, a goal-oriented tutorial designed to improve healthy nutrition and exercising (Horowitz, Shilts, & Townsend, 2004; Shilts, Townsend, & Horowitz, 2004). Based on Social Cognitive Theory, the original tutorial was effective for sixth and eighth graders when administered as a 10-hour classroom setting in-person curriculum (e.g., Shilts, Horowitz, & Townsend, 2009). The tutorial’s lessons included nutrition and exercise basics as well as information about energy balance, food labels, fast food, and advertising.

The adapted version shares the content of EatFit (i.e., all the facts covered in EatFit were also covered in GistFit), however, it also emphasizes the bottom-line meaning of nutrition and exercise according to research on fuzzy-trace theory (e.g., Reyna & Mills, 2014; Wolfe,

Reyna, Widmer, Cedillos, et al., 2015). The tutorial was adapted for college-aged students by removing repeated information or superficially relevant content (e.g., repeated information about nutrition label reading and calorie count) and by adapting class exercises for online use (e.g., class exercises were replaced with illustrative examples and key points). Factual information was updated from the original EatFit based on recent research, always in accordance with the principles guiding its design and content. In addition to presenting nutrition and exercise verbatim facts, the tutorial emphasized the essential decision-relevant meaning of information. Moreover, in GistFit, short bottom-line (gist) summaries of important points from each lesson were provided at the end of each lesson to facilitate the encoding and long-term retention of core information.

GistFit is an ITS built using SKO platform, which is a web-based system using artificial intelligence techniques to mimic one-on-one human tutoring. Throughout the tutorial, an avatar presents information orally and in writing, while simultaneously illustrating content with figures and images (e.g., examples of food containing each nutrient or a sequence of images demonstrating how to correctly perform aerobic exercises; Graesser, VanLehn, Rosé, Jordan, & Harter, 2001). A screen shot of GistFit from the learner's perspective can be found in Figure 1. Three female avatars (each of a different apparent ethnicity) deliver information throughout the tutorial. The tutorial is self-administered and took participants approximately 90 minutes to complete, although some engaged with the interactive elements of the tutorial for longer.

The control groups all received tutorials about the same unrelated health topic: genetic testing and breast cancer risk. Delivery format was the same for all conditions (i.e., all tutorials were constructed using the SKO platform) and time on task was the same as GistFit (i.e., approximately 90 minutes).

Measures embedded within the tutorial

Multiple-choice questions: Throughout the tutorial, participants were encouraged to think actively about content with the help of seven multiple-choice questions on the topics studied to check for understanding. Each multiple-choice question was presented at key points within the tutorial after participants studied the topic in question, such as “Is cholesterol ever good?” If the participant selected the wrong answer (i.e., “Yes... because you find it in egg yolks, which are healthy” which is wrong because egg yolks have bad cholesterol), then they had a chance to try one more time. Once they the answer (i.e., “Yes...because it helps clear your blood vessels” which referred to the good cholesterol), they received feedback and a reminder of differences between good and bad cholesterol before moving to the next lesson. Items were scored as correct or incorrect, and responses were averaged. Higher scores representing better gist understanding of the reasons (i.e., meaningful messages) behind the essential facts. In this case, if participants understand that some kinds of cholesterol are “good” because they help remove fat-like substances from the body, they will choose the option of “clearing the blood vessels.”

Analytic questions: Participants were also encouraged to actively generate and elaborate on explanations of content materials through analytic questions (i.e., questions that require a

student to not only know about concepts but apply those concepts to new situations; Arnott et al., 2008). Four questions were included throughout the tutorial: the reasons why participants would want to improve and maintain their health, the actions they could take to eat healthily, the actions they could take to incorporate exercise into their routines, and the aspects of their lives under their control that could affect their food and exercise choices. These questions encouraged participants to engage in tutorial dialogues that took advantage of a Latent Semantic Analysis (LSA) feature of the SKO platform. LSA enables the avatars to process participants' verbal input and respond to what people type in conversation (Graesser et al., 2000, 2004).

LSA compares conceptual and semantic similarities between text entered by participants and expectation texts prepared by experts in nutrition (i.e., list of ideal answers). The expectation texts included the core gist content of a good answer, such as “avoid becoming overweight” to answer the reasons to maintain and improve health. Because each person had a different answer, the tutorial responses were individualized depending on the level of relatedness between the input (e.g., “To avoid getting diseases in the future.”) and the expectation text (e.g., “lower risk of disease”). The tutorial would then respond according to how the participant was doing. If the participant was doing well, the tutorial encouraged them to continue on the same theme (e.g., “Doing well so far”). If they were off topic, the tutorial encouraged the participant to get back on track (e.g., “You might be getting off track”). If the participant was repeating information, the tutorial would point it out (e.g., “Right, but you said something similar before”). The ability to automatically tailor responses to each participant makes GistFit a unique obesity prevention tutorial (Lairson, Chan, Hang, Junco, & Vermon, 2011; Wolfe et al., 2013; Wolfe, Reyna, Widmer, Cedillos-Whynott, et al., 2015). To avoid cuing specific verbatim memories and instead increase recall of gist values retained by participants, we only provided general pumps (i.e., general encouragements) and not specific prompts (i.e., cues to say more about a specific topic; Brainerd & Reyna, 2002; Mills, Reyna, & Estrada, 2008).

To illustrate, the following selection was part of a dialogue between GistFit and a participant. Each response was judged to be appropriate and encouraged the participant to continue talking about her reasons to improve and maintain her health. Typos and spelling mistakes are included in these transcripts, demonstrating that GistFit performs robustly to such errors.

Participant: Improving health will increase my life expectancy, reduce my stress and anxiety, and allow me to be a happier person.

GistFit: Great job!

Participant: Improving my health will also allow me to be more active.

GistFit: Doing well so far!

In this example, GistFit recognized that the responses matched the expectation texts and encouraged the participant to continue writing. However, when the participant was asked to describe some of the aspects of her life that she could control that affected her food and exercise choices, GistFit recognized that a participant's response was good but repetitive.

Participant: The weather can affect my exercise choice, social situations can affect my eating habits.

GistFit: Doing well so far!

Participant: If I am upset or Happy I am more likely to eat more food.

GistFit: You're doing great! Keep thinking about specific factors that can affect your choices.

Participant: If I am depressed I will also eat more food, I have to learn to eat for hunger and not anxiety.

GistFit: Right, but you have said something similar before.

With each reply from GistFit, the avatar reports an overall coverage (CO) score ranging from 0 to 1, based on the extent to which the total text input matched the expectation text. To prevent mere repetition of words from superficially inflating the CO score, SKO has a semantic processing engine that evaluates the participant's response according to both relevance and novelty, resulting in four possible designations of old-relevant, new-relevant, old-irrelevant, and new-irrelevant (Wolfe et al., 2013). GistFit was programmed such that only new-relevant text would raise the CO score, while responses that fell in other categories would lower it. In order to establish the reliability of the tutorial CO scores of the semantic similarity of participants' answers, the score for each sentence was compared to expert judgment scores. Two independent judges blind to the tutorial's CO scores used the same rubric as the tutorial to rate each sentence. The two judges had .98 agreement. The mean score of the two judges was compared with the tutorial CO scores, resulting in a .91 agreement. The final CO score for each question (i.e., a measure of the degree to which the expectations have been met by all of the learner's responses, combined across all sentences entered by the participant) was used for analyses, and a higher value indicated more overall coverage of content compared to expectations texts.

In addition to CO score, participants were evaluated according to level of engagement in the dialogues with the tutorial (i.e., amount of interaction chosen by the respondent), categorized as either *none* (i.e., no response ever given to avatar), *some* (i.e., some avatar questions were answered) or *all* (i.e., every question was answered). This variable could be treated as a grouping factor with three levels. Alternatively, it could be combined with the control group to create a 4-point ordinal scale of engagement in tutorial dialogues (no tutorial, none, some, or all). A higher score on both scales implied a more active experience.

Survey measures

Knowledge: Knowledge was assessed in two ways. The first scale was a measure of verbatim knowledge and included 22 multiple-choice questions assessing specific, rote content about nutrition, use of the nutrition labels, and exercise ($\alpha = 0.65$). For example, participants answered questions such as "How many repetitions of body weight exercises should you do to really work your muscles?" from four options (i.e., "Several sets of 10–15," "Count to your age three times," "Several sets of 5," or "Several sets of 30"—the first option being the correct answer). The second scale was a measure of gist knowledge, which

required participants to transfer knowledge from specific information presented in the tutorial to another application as a way of demonstrating general, meaningful understanding about the topic. Participants provided answers to 22 multiple-choice questions assessing knowledge about nutrition, use of nutrition label, and exercise ($\alpha = 0.68$). For example, participants were prompted to answer “What is a healthy reason to choose foods high in calcium?” with one of the following options: “Calcium-rich foods allow your cells to repair your bones” (the correct answer), “A high calcium intake is needed to properly absorb Vitamin D, a nutrient important for healthy teeth,” and “Calcium-rich foods usually taste good.” In both scales, items were rated as correct or incorrect, and responses were averaged, with higher scores representing better knowledge. It should be noted that knowledge tests contain different heterogeneous items and it is not necessarily to be expected that they would all measure one unitary concept. However, these knowledge items were extracted from published nutritional and lifestyle scientific literature and discussed with a nutrition expert.

Gist comprehension: A measure of gist comprehension was constructed to assess understanding and integrating of tutorial information about nutrition and exercise ($\alpha = 0.87$). The scale included 22 items such as “To lose weight, I should make consistent changes to my eating and exercise habits because most fad diets are unsafe or ineffective.” Items were rated on a 6-point Likert scale from *strongly disagree* to *strongly agree*, and responses were averaged. Higher scores imply better gist comprehension.

Gist principles: Gist measures contained 57 general simple healthy values and principles that were organized according to three subscales: *lifestyle* (21 items, $\alpha = 0.95$; e.g., “Better to eat food low in sugar now than to deal with health consequences later”), *nutrition* (20 items, $\alpha = 0.92$; e.g., “Any amount of trans-fat is not good for you”), and *exercise* (16 items, $\alpha = 0.87$; e.g., “Avoid watching TV for long periods of time”). Participants could endorse each item on a 5-point Likert scale from *strongly disagree* to *strongly agree*. Responses were averaged and higher scores indicated a greater tendency to endorse gist-based healthy values.

Behavioral intentions: A measure of behavioral intentions of healthy nutrition and exercising was constructed from Baker, Little, and Brownell’s (2003) ratings to 8 Likert-type items ($\alpha = 0.89$), such as “I plan to be a healthy eater” and “I plan to be physically active.” The ratings were measured on a 5-point scale from *strongly disagree* to *strongly agree* and averaged. Higher scores implied greater intention to perform healthier behavior.

Behavioral measures: Self-reported behavioral measures were taken from the American Heart Association’s (AHA) national goals for cardiovascular health promotion and disease reduction (Lloyd-Jones et al., 2010), including measures of healthy nutrition (amount of fruit, vegetables, fish, whole grain, sugar and sodium ingestion per week) and exercise (number of hours of moderate and vigorous activity per week).

Procedure

The tutorials were administered online by undergraduate or graduate research assistants in a controlled laboratory environment. Participants were informed prior to enrollment that they

would be randomized to learn about one of two topics (all control conditions were about the same topic). After the tutorial, participants answered a survey with the aforementioned questions.

Data Analyses

Data were analyzed using IBM SPSS Statistics (Version 21.0, IBM Corp., Armonk, NY). First, we ran bivariate correlations to examine the extent to which measures of knowledge, gist comprehension, and gist principles, behavioral intentions, and self-reported behavior shared variance regarding healthy nutrition and exercise. Next, we ran bivariate correlations between survey measures and the measures embedded within the tutorial—multiple-choice questions, final CO score, and level of engagement in tutorial dialogue—to test associations between active learning and increases in knowledge, gist comprehension, and gist principles.

Then, to investigate tutorial dialogues and how they were related to tutorial outcomes, we conducted a 4 (engagement) \times 2 (research site) analysis of variance for each of the dependent measures. The factor *engagement* included the four levels described above, separating participants according to who engaged in *all* tutorial dialogues (N = 19), *some* of them (N = 14), *none* of them (N = 12), or those who received the *control* tutorial (N = 183). The goal was to determine whether different levels of engaging in dialogues was associated with better learning outcomes after interaction with an ITS.

Results and Discussion

Correlations of Knowledge, Comprehension, Principles, Intentions and Behavior

Preliminary to discussing engagement, we begin with self-reported behavior and behavioral intentions. Behavioral intentions have been found to be a predictor of behavior (for a review, see Greaves et al., 2011; Webb & Sheeran, 2006). Consistent with that literature, behavioral intentions correlated with self-reported healthy nutrition and exercising (Table 2). In addition, healthier self-reported nutrition and exercising were related to greater verbatim knowledge (Table 2), which included specific tutorial content.

Self-reported nutrition and overall behavioral intentions were also related to endorsements of gist principles and to gist comprehension (Table 2). Gist principles of overall lifestyle correlated with both healthy nutrition and exercise (Table 2). Endorsement of healthy values—expressed as simple gist principles—would be expected to correlate with healthier behavior and intentions, but this is the first demonstration of these relationships to gist in the domain of obesity prevention (e.g., Mills et al., 2008; Reyna, Estrada, DeMarinis, Myers, Stanisiz, & Mills, 2011). Thus, these results for nutrition and exercise extend prior research on psychosocial mediators of risky (i.e., unhealthy) behaviors to new behavioral domains.

Endorsements of gist principles of lifestyle, nutrition, and exercise were associated with higher verbatim and gist knowledge scores as well as gist comprehension (Table 2). Among the latter variables, gist comprehension had the strongest relationships with endorsements of gist principles and with overall behavioral intentions, compared to either verbatim or gist knowledge. These results suggest that deeper levels of understanding, rather than merely

rote knowledge of details, may be important in establishing commitments to healthy values and intentions, which is consistent with prior findings.

Accuracy and Level of Tutorial Engagement

The GistFit participants who produced more correct answers to multiple-choice questions to check understanding embedded within the tutorial also showed better performance in gist knowledge [$r(32) = .55, p = .001$] and gist comprehension [$r(32) = .45, p = .01$]. The average correct response to multiple-choice questions was .78 ($SD = .27$). CO scores were unrelated to knowledge and comprehension (most likely due to restriction of range: approximately 60% of the answers were between .50 and .70).

Gist knowledge, gist comprehension, and verbatim knowledge were correlated with the ordinal measure of active engagement in tutorial dialogues (Table 2). The higher the level of engagement from participants who were exposed to GistFit (using a three-point ordinal scale of engagement), the more likely they were to perform better in factual questions as well as those that require transfer of knowledge (Table 2). Actively taking the tutorial in the sense of engaging in dialogues was associated with better learning and gist understanding. These relationships were confined to knowledge measures when the control group was included in the ordinal measure of engagement (i.e., control group, and *none*, *some*, or *all* questions answered). Taken together, these findings are consistent with fuzzy-trace theory's prediction that the ability to extract gist, and the application of that ability in engaging in active reasoning (interacting with the tutorial), would improve retention and transfer (Reyna, 2013; Reyna, Chapman, Dougherty, & Confrey, 2012).

Further analyses compared groups with differing levels of engagement with the tutorial separately for dependent measures of knowledge and comprehension. Table 3 displays the means and standard deviation for each of the levels of engagement. The ANOVAs on Gistfit alone (three levels of engagement) showed significant differences for verbatim knowledge, gist knowledge, and gist comprehension, respectively [$F(2, 39) = 4.27, p = .025, MSE = 0.09, \eta_p^2 = .17$; $F(2, 39) = 5.96, p = .006, MSE = 0.13, \eta_p^2 = .23$; and $F(2, 39) = 5.70, p = .007, MSE = 2.99, \eta_p^2 = .23$]. Including the control group produced the same three significant results for verbatim knowledge, gist knowledge, and gist comprehension: $F(3, 243) = 9.06, p < .001, MSE = 0.18, \eta_p^2 = .10$; $F(3, 243) = 4.46, p = .005, MSE = 0.11, \eta_p^2 = .05$; and $F(3, 243) = 3.86, p = .01, MSE = 2.11, \eta_p^2 = .05$. Consistent with our predictions, actively learning the gist of nutrition and exercise by engaging in *all* or *some* dialogues with the tutorial was associated with greater verbatim knowledge, compared to not engaging or being in the control group (i.e., not taking the Gistfit tutorial). Table 3 also shows that similar significant differences were observed for gist comprehension. In addition, engaging in *all* dialogues was associated with greater gist knowledge compared to both answering *no* questions or the control group (and *some* was greater than *none*).

Thus, these results suggest that actively engaging in dialogues with the tutorial was related to the extent to which participants remembered and comprehended the gist of information after the tutorial. Note that these were not randomly assigned groups, but the level of engagement was self-selected, and, therefore, causal conclusions cannot be made about engagement. However, we cannot dismiss the results of other studies showing that

encouraging active learning and giving feedback stimulated students to extract gist understanding and construct their knowledge (proactive thinking; e.g., Chi, 2009; VanLehn et al., 2007; Wolfe et al., 2013; Wolfe, Reyna, Widmer, Cedillos, et al., 2015). One possible kind of self-selection might be that people who engage more in dialogues with the tutorial are smarter to begin with (or that smart people may engage more) and, thus, intelligence might account for differences in outcomes (e.g., Ackerman, Kanfer, & Goff, 1995; Chamorro-Premuzic, Furnham, & Ackerman, 2006). In order to test this hypothesis, we used differences in numeracy as a rough proxy for intelligence (Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012; Liberali, Reyna, Furlan, Stein, & Pardo, 2012). Analyses showed those who differed in engagement with the tutorial did not differ significantly in numeracy ($F_s < 1$) using the 15-item objective numeracy scale (Lipkus, Samsa, & Rimer, 2001; Peters, Dieckmann, Dixon, Hibbard, & Mertz, 2007). Hence, a priori differences in intelligence did not seem to account for differences in outcomes that are associated with engagement with the tutorial.

Conclusions

Engaging in tutorial dialogues was associated with increases in not only knowledge, but also transfer of that knowledge through understanding the behaviors that prevent obesity, as assessed in the gist comprehension test. Throughout the obesity prevention tutorial, artificial intelligence was used to encourage participants to reason actively about the topics they learned and to extract gist understanding. According to fuzzy-trace theory, gist understanding supports productive thought that involves the transfer of knowledge that is essential for behavior change (i.e., the ability to apply prior learning to novel instances; Wolfe, Reyna, & Brainerd, 2005). The GistFit ITS was designed to facilitate connecting the dots among learned facts and making meaningful inferences that go beyond surface (rote) information learned (Lloyd & Reyna, 2009; Wolfe et al., 2013; Wolfe, Reyna, Widmer, Cedillos, et al., 2015). Those who responded more to the ITS by engaging in dialogue with the avatar seem to have found the most benefit in transfer of information, which is consistent with our theoretical predictions.

Factual knowledge is a target of both traditional and fuzzy-trace theory interventions because it is necessary for behavior change; however, it is not sufficient (e.g., Reyna & Mills, 2014; Wolfe, Reyna, Widmer, Cedillos, et al., 2015). In GistFit, individuals learned both verbatim facts and bottom-line (gist) meaning of those facts and interactions with the tutor were related to increases in knowledge and understanding. Limitations inherent in correlational analyses should be acknowledged. Correlations between measures that are taken at the same time point might share some method variance. However, method variance cannot be the whole explanation, because not all of the measures with the same response scale correlated with each other (e.g., behavioral intentions and gist comprehension). In addition, there is independent evidence that these measures capture substantive variance beyond method variance because they do not just correlate with each other and not predict other things (e.g., Mills et al., 2008; Reyna et al., 2011; Reyna & Mills, 2014; Wolfe et al., 2015).

The fact that the GistFit outperformed the control group on these questions does not demonstrate that the gist enhancements drove the effects. However, other studies have demonstrated that gist-based reasoning improves knowledge and decision-making, using random assignment to compare verbatim and gist versions of curricula with similar content (e.g., Reyna & Mills, 2014; Wolfe, Reyna, Widmer, Cedillos, et al., 2015). Consistent with this interpretation, higher scores on gist knowledge questions that require transfer of knowledge (i.e., extrapolating concepts to novel questions) seem to be a test of deeper understanding from active learning. Verbatim knowledge was also positively correlated with GistFit participants who interacted more with the tutorial, which is consistent with research that has shown that testing immediately after an event reinforces verbatim memory, in particular after previous testing like the multiple-choice questions embedded within the tutorial (Brainerd & Reyna, 1996, 2002). Nevertheless, merely memorized verbatim facts decay more rapidly over time (e.g., Brainerd et al., 2009), and do not have as strong an influence on behavior change (e.g., Reyna & Mills, 2014).

Another limitation of the present study is that level of engagement was self-selected by participants, and, thus, the significant differences may be due to causes other than engagement. However, other studies have shown that active engagement was causally related to increased knowledge (Chi, 2009; Wolfe, Britt, Petrovic, Albrecht, & Kopp, 2009; see also Widmer, Wolfe, Reyna, Cedillos-Whynott, Brust-Renck, & Weil, 2015). Although participants might come to a study with different levels of engagement, and that, in combination with the content of the tutorial, might be important in establishing commitments to healthy values and intentions. Nevertheless, as with any correlational study it is possible that some other, unmeasured variable may explain the results (e.g., motivation). Future research should measure pre-existing motivation separately from tutorial engagement and also randomly assign participants to different levels of engagement in GistFit to test this hypothesis that people who are more responsive to the intelligent tutor's questions perform better in posttest outcomes. Another limitation is that the design included only a posttest without a pretest. Random assignment of participants to the different tutorials probably adequately eliminated baseline differences, but a pretest would have permitted a more detailed analysis of associations between engagement and learning gains. Another argument against prior differences is that the participants did not differ in demographics, numeracy, or risk factors as measured by American Heart Association cardiovascular health measures. Further research to confirm the effectiveness of this sort of tutorial could include assessing the effects of engaging in interactions with the tutorial on behavioral changes over time; we expect that the more people engage in dialogues, the more they maintain changes over time. Although previous studies have shown that emphasizing the bottom-line gist of each lesson was effective at reducing risk taking over time (e.g., Reyna & Mills, 2014), no study to our knowledge has compared the interactive effects of a fuzzy-trace theory-based ITS over time.

In sum, although the nutrition and exercise industry bombards the public with commercial messages, there are far too few theoretically motivated approaches. There are even fewer approaches that combine artificial intelligence with the latest cognitive theories. A major issue with the use of artificial intelligence for health communication has been the difficulty of emulating human-like interaction. As illustrated in the current study, emerging discourse technologies can be used effectively to encourage participants to engage in interactive

behaviors as they construct their knowledge (Chi, 2009; VanLehn, 2011). When the use of technology is guided by a sound theoretical understanding, effective communication is more likely (e.g., Wolfe, Reyna, Widmer, Cedillos, et al., 2015; Wolfe, Reyna, Widmer, Cedillos-Whynott, et al., 2015). The current ITS was directed at obesity prevention by encouraging productive thinking and extraction of gist-based understanding through interactions with the tutor. Results of this study are consistent with beneficial effects of such interaction on gist understanding, an “active ingredient” in interventions to promote behavior change, and provide a motivation for further research using random assignment to levels of engagement with artificial tutors.

Acknowledgments

Preparation of this manuscript was supported in part by the National Cancer Institute of the National Institutes of Health under Award Number R21CA149796 to the second and fourth authors. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health. We gratefully acknowledge the assistance of Christopher R. Fisher, Audrey Weil, Emily Lopes, Amrita Rao, Suveera Dang, Sharjeel Chaudhry, Xuan Zhang, Edward Shin, and Lindsay Dower. The authors thank consulting SKO creator Dr. Xiangen Hu, PhD (University of Memphis).

References

- ABGC—American Board of Genetic Counseling, Inc. About ABGC [Web page]. Retrieved from http://www.abgc.net/About_ABGC/GeneticCounselors.asp
- Ackerman PL, Kanfer R, Goff M. Cognitive and noncognitive determinants and consequences of complex skill acquisition. *Journal of Experimental Psychology: Applied*. 1995; 1(4):270. doi: 10.1037/1076-898X.1.4.270
- Ajzen I. The theory of planned behaviour: Reactions and reflections. *Psychology & Health*. 2011; 26(9):1113–1127. DOI: 10.1080/08870446.2011.613995 [PubMed: 21929476]
- Arnott E, Hastings P, Allbritton D. Research Methods Tutor: Evaluation of a dialogue-based tutoring system in the classroom. *Behavior Research Methods*. 2008; 40(3):694–698. DOI: 10.3758/BRM.40.3.694 [PubMed: 18697663]
- Baker CW, Little TD, Brownell KD. Predicting adolescent eating and activity behaviors: The role of social norms and personal agency. *Health Psychology*. 2003; 22(2):189–198. DOI: 10.1037/0278-6133.22.2.189 [PubMed: 12683739]
- Betsch C, Brewer NT, Brocard P, Davies P, Gaissmaier W, Haase N, Leask J, Renkewitz F, Renner B, Reyna VF, Rossmann C, Sachse K, Schachinger A, Siegrist M, Stryk M. Opportunities and challenges of web 2.0 for vaccination decisions. *Vaccine*. 2012; 28(30):3727–3733. DOI: 10.1016/j.vaccine.2012.02.025
- Brainerd CJ, Reyna VF. Mere memory testing creates false memories in children. *Developmental Psychology*. 1996; 32:467–478. DOI: 10.1037/0012-1649.32.3.467
- Brainerd CJ, Reyna VF. Fuzzy-trace theory and false memory. *Current Directions in Psychological Science*. 2002; 11:164–168. DOI: 10.1111/1467-8721.00192
- Brainerd CJ, Reyna VF, Howe ML. Trichotomous processes in early memory development, aging, and cognitive impairment: A Unified Theory. *Psychological Review*. 2009; 116:783–832. DOI: 10.1037/a0016963 [PubMed: 19839684]
- Brust-Renck, PG., Reyna, VF., Wolfe, CR., Cedillos, EM., Widmer, CL., Fisher, CR., Wilhelms, EA., Chaudhry, S., Lopes, EA., Wampler, AT., Dang, S., Nollet, ZW. Randomized control trial of an obesity-prevention curriculum to improve psychosocial mediators to health outcomes (based on fuzzy-trace theory). Poster presented at the 33rd Annual Meeting of the Society for Judgment and Decision Making, Minneapolis; MN. 2012.
- Brust-Renck PG, Royer CE, Reyna VF. Communicating numerical risk: Human factors that aid understanding in health care. *Reviews of Human Factors and Ergonomics*. 2013; 8(1):235–276. DOI: 10.1177/1557234X13492980

- Chi MTH. Active-Constructive-Interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*. 2009; 1:73–105. [PubMed: 25164801]
- Chi MTH, Siler SA, Jeon H, Yamauchi T, Hausmann RG. Learning from human tutoring. *Cognitive Science*. 2005; 25(4):471–533. doi: 1007/s15516709cog2504_1.
- Clariana RB, Koul R. Multiple-try feedback and higher-order learning outcomes. *International Journal of Instructional Media*. 2005; 32(3):239–245.
- Clariana RB, Koul R. The effects of different forms of feedback on fuzzy and verbatim memory of science principles. *British Journal of Educational Psychology*. 2006; 76:259–270. DOI: 10.1348/000709905X39134 [PubMed: 16719963]
- Chamorro-Premuzic T, Furnham A, Ackerman PL. Incremental validity of the Typical Intellectual Engagement Scale as predictor of different academic performance measures. *Journal of Personality Assessment*. 2006; 87:261–268. DOI: 10.1207/s15327752jpa8703_07 [PubMed: 17134334]
- Cokely ET, Galesic M, Schulz E, Ghazal S, Garcia-Retamero R. Measuring risk literacy: The Berlin Numeracy Test. *Judgment and Decision Making*. 2012; 7:25–47. Retrieved from: <http://journal.sjdm.org/11/11808/jdm11808.html>.
- Cooper Z, Helen HA, Hawker DM, Byrne S, Bonner G, Eeley E, Fairburn CG. Testing a new cognitive behavioural treatment for obesity: A randomized controlled trial with three-year follow-up. *Behaviour Research and Therapy*. 2010; 48(8):706–713. DOI: 10.1016/j.brat.2010.03.008 [PubMed: 20691328]
- Downs JS, Bruin de Bruine WD, Fischhoff B. Parents' vaccination comprehension and decisions. *Vaccine*. 2008; 26:1595–607. DOI: 10.1016/j.vaccine.2008.01.011 [PubMed: 18295940]
- Fernbach PM, Rogers T, Fox CR, Sloman SA. Political extremism is supported by an illusion of understanding. *Psychological Science*. 2013; 24(6):939–946. [PubMed: 23620547]
- Fishbein M. A reasoned action approach to health promotion. *Medical Decision-making*. 2008; 28(6): 834–844. DOI: 10.1177/0272989X08326092 [PubMed: 19015289]
- Graesser AC, Lu S, Jackson GT, Mitchell HH, Ventura M, Olney A, Louwerse MM. AutoTutor: A tutor with dialogue in natural language. *Behavior Research Methods, Instruments, & Computers*. 2004; 36(2):180–192.
- Graesser AC, McNamara DS, VanLehn K. Scaffolding deep comprehension strategies through Point & Query, AutoTutor, and iSTART. *Educational Psychologist*. 2005; 40:225–234.
- Graesser AC, VanLehn K, Rosé CP, Jordan PW, Harter D. Intelligent tutoring systems with conversational dialogue. *AI Magazine*. 2001; 22:39.
- Graesser AC, Wiemer-Hastings P, Wiemer-Hastings K, Harter D, Tutoring Research Group, T. R. G. & Person N. Using latent semantic analysis to evaluate the contributions of students in AutoTutor. *Interactive Learning Environments*. 2000; 8(2):129–147. DOI: 10.1076/1049-4820
- Greaves CJ, Sheppard KE, Abraham C, Hardeman W, Roden M, Evans PH, Schwarz P, The IMAGE Study Group. Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions. *BMC Public Health*. 2011; 11(119)doi: 10.1186/1471-2458-11-119
- Horowitz M, Shilts MK, Townsend MS. EatFit: A goal-oriented intervention that challenges adolescents to improve their eating and fitness choices. *Journal of Nutrition Education and Behavior*. 2004; 36(1)(06):43–44. 60128–0. DOI: 10.1016/S1499-4046 [PubMed: 14756982]
- Hu, X., Cai, Z., Han, L., Craig, SD., Wang, T., Graesser, AC. Proceedings of the 2009 Conference of Artificial Intelligence in Education: Building Learning Systems That Care: From Knowledge Representation to Affective Modeling. Amsterdam, The Netherlands: IOS Press; 2009. AutoTutor Lite; p. 802
- Hu, X., Han, L., Cai, Z. Paper presented to the Society for Computers in Psychology. Chicago, IL: 2008. Semantic decomposition of student's contributions: an implementation of LCC in AutoTutor Lite.
- Lairson DR, Chan W, Chang YC, Junco DJ, Vernon SW. Cost-effectiveness of targeted vs. tailored interventions to promote mammography screening among women military veterans in the United States. *Evaluation and Program Planning*. 2011; 34:97–104. DOI: 10.1016/j.evalprogplan.2010.07.003 [PubMed: 20810168]

- Li M, Chapman GB. Nudge to health: Harnessing decision research to promote health behavior. *Social and Personality Psychology Compass*. 2013; 7(3):187–198. DOI: 10.1111/spc3.12019
- Liberali JM, Reyna VF, Furlan S, Stein LM, Pardo ST. Individual differences in numeracy and implications for biases and fallacies in probability judgment. *Journal of Behavioral Decision Making*. 2011; 2:361–381. DOI: 10.1002/bdm.752
- Lipkus IM, Samsa G, Rimer BK. General performance on a numeracy scale among highly educated samples. *Medical Decision-making*. 2001; 21:37–44. DOI: 10.1177/0272989X0102100105 [PubMed: 11206945]
- Lloyd FJ, Reyna VF. A web exercise in evidence-based medicine using cognitive theory. *Journal of General Internal Medicine*. 2001; 16(2):94–99. DOI: 10.1111/j.1525-1497.2001.00214.x [PubMed: 11251760]
- Lloyd, FJ., Reyna, VF. Clinical gist and medical education: Connecting the dots. Vol. 302. *Journal of the American Medical Association*; 2009. p. 1332-1333.
- Lloyd-Jones D, Adams RJ, Brown TM, Carnethon M, Dai S, De Simone G, Wylie-Rosett J. Heart Disease and Stroke Statistics—2010 Update: A Report From the American Heart Association. *Circulation*. 2010; 121:e46–e215. [PubMed: 20019324]
- Mills B, Reyna VF, Estrada S. Explaining contradictory relations between risk perception and risk taking. *Psychological Science*. 2008; 19(5):429–433. DOI: 10.1111/j.1467-9280.2008.02104.x [PubMed: 18466401]
- Ogden CL, Carroll MD, Kit BK, Flegal KM. Prevalence of childhood and adult obesity in the United States, 2011–2012. *JAMA*. 2014; 311(8):806–814. DOI: 10.1001/jama.2014.732 [PubMed: 24570244]
- Peters E. A perspective on eating behaviors from the field of judgment and decision making. *Annals of Behavioral Medicine*. 2009; 38(Suppl 1):S81–S87. DOI: 10.1007/s12160-009-9121-8 [PubMed: 19787307]
- Peters E, Dieckmann N, Dixon A, Hibbard JH, Mertz CK. Less is more in presenting quality information to consumers. *Medical Care Research and Review*. 2007; 64(2):169–190. DOI: 10.1177/1077558706298290 [PubMed: 17406019]
- Reyna VF. A new intuitionism: Meaning, memory, and development in fuzzy-trace theory. *Judgment and Decision-making*. 2012a; 7(3):332–359. Retrieved from journal.sjdm.org/11/111031/jdm111031.html. [PubMed: 25530822]
- Reyna VF. Risk perception and communication in vaccination decisions: A fuzzy-trace theory approach. *Vaccine*. 2012b; 30(25):3790–3797. DOI: 10.1016/j.vaccine.2011.11.070 [PubMed: 22133507]
- Reyna, VF. Intuition, reasoning, and development: A fuzzy-trace theory approach. In: Barrouillet, P., Gauffroy, C., editors. *The development of thinking and reasoning*. Hove, UK: Psychology Press; 2013. p. 193-220.
- Reyna VF, Adam MB. Fuzzy-trace theory, risk communication, and product labeling in sexually transmitted diseases. *Risk Analysis*. 2003; 23(2):325–342. DOI: 10.1111/1539-6924.00332 [PubMed: 12731817]
- Reyna, VF., Brainerd, CJ. A fuzzy-trace theory of reasoning and remembering: Paradoxes, patterns, and parallelism. In: Healy, A.Kosslyn, S., Shiffrin, R., editors. *From learning processes to cognitive processes: Essays in honor of William K Estes*. Vol. 2. Hillsdale, NJ: Erlbaum; 1992. p. 235-259.
- Reyna VF, Brainerd CJ. Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences*. 1995; 7(95):1–75. 90031–4. DOI: 10.1016/1041-6080
- Reyna VF, Brainerd CJ. Dual processes in decision-making and developmental neuroscience: A fuzzy-trace model. *Developmental Review*. 2011; 31:180–206. DOI: 10.1016/j.dr.2011.07.004 [PubMed: 22096268]
- Reyna, VF., Chapman, SB., Dougherty, M., Confrey, J. *The adolescent brain: Learning, reasoning and decision-making*. Washington DC: American Psychological Association; 2012.
- Reyna VF, Estrada SM, DeMarinis JA, Myers RM, Stanisz JM, Mills BA. Neurobiological and memory models of risky decision-making in adolescents versus young adults. *Journal of*

- Experimental Psychology: Learning, Memory, and Cognition. 2011; 37(5):1125–1142. DOI: 10.1037/a0023943
- Reyna VF, Farley F. Risk and rationality in adolescent decision-making: Implications for theory, practice, and public policy. *Psychological Science in the Public Interest*. 2006; 7(1):1–44. DOI: 10.1111/j.1529-1006.2006.00026.x [PubMed: 26158695]
- Reyna VF, Lloyd FJ. Physician decision-making and cardiac risk: Effects of knowledge, risk perception, risk tolerance, and fuzzy processing. *Journal of Experimental Psychology*. 2006; 12(3): 179–195. DOI: 10.1037/1076-898X.12.3.179 [PubMed: 16953744]
- Reyna, VF., Lloyd, FJ., Brainerd, CJ. Memory, development, and rationality: An integrative theory of judgment and decision-making. In: Schneider, S., Shanteau, J., editors. *Emerging perspectives on judgment and decision research*. New York: Cambridge University Press; 2003. p. 201-245.
- Reyna VF, Mills BA. Theoretically motivated interventions for reducing sexual risk taking in adolescence: A randomized controlled experiment applying fuzzy-trace theory. *Journal of Experimental Psychology: General*. 2014; 143(4):1627–1648. DOI: 10.1037/a0036717 [PubMed: 24773191]
- Shilts MK, Horowitz M, Townsend MS. Guided goal setting: Effectiveness in a dietary and physical activity intervention with low-income adolescents. *International Journal of Adolescent Medicine and Health*. 2009; 21(1):111–122. [PubMed: 19526701]
- Shilts MK, Townsend MS, Horowitz M. An innovative approach to goal setting for adolescents: Guided goal setting. *Journal of Nutrition Education and Behavior*. 2004; 36(3):155–156. DOI: 10.1016/S1499-4046(06)60153-X [PubMed: 15202992]
- VanLehn K. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*. 2011; 46(4):197–221. DOI: 10.1080/00461520.2011.611369
- VanLehn K, Graesser AC, Jackson GT, Jordan P, Olney A, Rosé CP. When are tutorial dialogues more effective than reading? *Cognitive Science*. 2007; 31(1):3–62. DOI: 10.1080/03640210709336984 [PubMed: 21635287]
- Webb TL, Sheeran P. Does changing behavioral intentions engender behavioral change? A meta-analysis of the experimental evidence. *Psychological Bulletin*. 2006; 132(2):249–268. DOI: 10.1037/0033-2909.132.2.249 [PubMed: 16536643]
- Wertheimer, M. *Productive thinking*. New York: Harper; 1959.
- Wertheimer, M. A Gestalt perspective on the psychology of thinking. In: Glatzeder, BM, Goel, V., von Muller, A., editors. *Towards a theory of thinking: Building blocks for a conceptual framework*. Heidelberg: Springer; 2010. p. 49-58.
- Widmer CL, Wolfe CR, Reyna VF, Cedillos-Whynott EM, Brust-Renck PG, Weil AM. Tutorial dialogues and gist explanations of genetic breast cancer risk. *Behavior Research Methods*. 2015; 47:632–648. DOI: 10.3758/s13428-015-0592-1 [PubMed: 25921818]
- Wolfe CR, Britt MA, Petrovic M, Albrecht M, Kopp K. The efficacy of a web-based counterargument tutor. *Behavior Research Methods*. 2009; 41:691–698. DOI: 10.3758/BRM.41.3.691 [PubMed: 19587180]
- Wolfe CR, Fisher CR, Reyna VF, Hu X. Improving internal consistency in conditional probability estimation with an Intelligent Tutoring System and web-based tutorials. *International Journal of Internet Science*. 2012; 7:38–54.
- Wolfe CR, Reyna VF. Assessing semantic coherence and logical fallacies in joint probability estimates. *Behavior Research Methods*. 2010a; 42(2):366–372. DOI: 10.3758/BRM.42.2.373 [PubMed: 20479168]
- Wolfe CR, Reyna VF. Semantic coherence and fallacies in estimating joint probabilities. *Journal of Behavioral Decision Making*. 2010b; 23(2):203–223. DOI: 10.1002/bdm.650
- Wolfe, CR., Reyna, VF., Brainerd, CJ. Fuzzy-trace theory: Implications for transfer in teaching and learning. In: Mestre, JP., editor. *Transfer of learning from a modern multidisciplinary perspective*. Greenwich, CT: Information Age Publishing; 2005. p. 53-88.
- Wolfe CR, Reyna VF, Widmer CL, Cedillos EM, Fisher CR, Brust-Renck PG, Weil AM. Efficacy of a web-based Intelligent Tutoring System for communicating genetic risk of breast cancer: A Fuzzy-

Trace Theory approach. *Medical Decision-making*. 2015; 35(1):46–59. DOI: 10.1177/0272989X14535983 [PubMed: 24829276]

Wolfe CR, Reyna VF, Widmer CL, Cedillos-Whynott EM, Brust-Renck PG, Weil AM, Hu X. Understanding genetic breast cancer risk: Processing loci of the BRCA Gist intelligent tutoring system. Manuscript submitted for publication. 2015

Wolfe CR, Widmer CL, Reyna VF, Hu X, Cedillos EM, Fisher CR, Brust-Renck PG, Williams TC, Damas I, Weil AM. The development and analysis of tutorial dialogues in AutoTutor Lite. *Behavior Research Methods*. 2013; 45:623–636. DOI: 10.3758/s13428-013-0352-z [PubMed: 23709166]



5/14

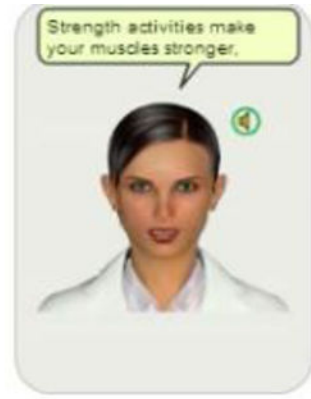


Figure 1.
A screen shot of GistFit.

Table 1

Percent Participant Characteristics by Condition

	GistFit	Control	Differences
Location (% Cornell)	64.4	64.1	$\chi^2(1) = .002, p = 0.963$
Age (mean/SD)	18.96 (1.17)	19.23 (1.77)	$t(241) = 1, p = 0.318$
Ethnicity (% Hispanic)	11.4	8.80	$\chi^2(1) = .29; p = 0.591$
Race (% White)	57.8	62.2	$\chi^2(1) = .31; p = 0.579$
Healthy Nutrition Behavior (mean/SD)	2.51 (1.32)	2.55 (1.18)	$t(247) = .22, p = 0.830$
Healthy Exercising Behavior (mean/SD)	2.54 (2.52)	2.53 (3.14)	$t(246) = -.03, p = 0.975$

Note. All irrelevant-content control tutorials were combined for analyses because differences among them were non-significant. Nutrition Behavior was based on the amount of healthy food ingested per week and Exercising Behavior was based on the number of hours of moderate and vigorous activity per week. SD = Standard Deviation.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2
Pearson Correlation Coefficients Between all Independent and Dependent Variables

	1	2	3	4	5	6	7	8	9	10	11
1 Level of Engagement (GistFit only)	1										
2 Level of Engagement (GistFit and control)	.97**	1									
3 Verbatim Knowledge	.31*	.27**	1								
4 Gist Knowledge	.37*	.16*	.57**	1							
5 Gist Comprehension	.35*	.12 [†]	.52**	.53**	1						
6 Gist Principles of Nutrition	.07	.09	.24**	.20**	.49**	1					
7 Gist Principles of Exercising	.06	.12 [†]	.14*	.15*	.35**	.63**	1				
8 Gist Principles of Lifestyle	.12	.10	.25**	.28**	.53**	.73**	.63**	1			
9 Behavioral Intentions	.03	.06	.17**	.07	.36**	.38**	.36**	.44**	1		
10 Healthy Nutrition Behavior	-.10	-.03	.12 [†]	.09	.12 [†]	.23**	.23**	.21**	.33**	1	
11 Healthy Exercising Behavior	.17	-.02	.03	.09	-.05	.09	.03	.14*	.19**	.18**	1

Note.

[†] $p = .05$.

* $p < .05$.

** $p < .01$.

Table 3

Means and Standard Deviations of Measures

	Min	Max	Control		GistFit Total		GistFit (All)		GistFit (Some)		GistFit (None)	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Verbatim Knowledge	0	1	0.64	0.15	0.74	0.16	0.78	0.13	0.74	0.16	0.66	0.19
Gist Knowledge	0	1	0.65	0.17	0.70	0.17	0.77	0.14	0.69	0.17	0.61	0.18
Gist Comprehension	0	5	4.48	0.76	4.61	0.80	4.89	0.49	4.60	0.80	4.19	1.06
Gist Principles of Nutrition	0	4	3.29	0.49	3.40	0.48	3.37	0.35	3.58	0.46	3.25	0.64
Gist Principles of Exercising	0	4	3.04	0.50	3.19	0.49	3.16	0.51	3.36	0.47	3.05	0.48
Gist Principles of Lifestyle	0	4	3.36	0.51	3.48	0.53	3.49	0.45	3.61	0.43	3.31	0.72
Behavioral Intentions	0	7	3.75	0.71	3.86	0.65	3.78	0.66	4.13	0.49	3.68	0.75

Note. All irrelevant-content control tutorials were combined for analyses because differences among them were non-significant. Nutrition Behavior was based on the amount of healthy food ingested per week and Exercising Behavior was based on the number of hours of moderate and vigorous activity per week. SD = Standard Deviation.