



HHS Public Access

Author manuscript

Learn Individ Differ. Author manuscript; available in PMC 2017 July 01.

Published in final edited form as:

Learn Individ Differ. 2016 July ; 49: 178–189. doi:10.1016/j.lindif.2016.06.009.

Understanding Genetic Breast Cancer Risk: Processing Loci of the *BRCA Gist* Intelligent Tutoring System

Christopher R. Wolfe¹, Valerie F. Reyna², Colin L. Widmer¹, Elizabeth M. Cedillos-Whynott¹, Priscila G Brust-Renck², Audrey M. Weil¹, and Xiangen Hu³

¹Miami University, Oxford, OH

²Cornell University, Ithaca, NY

³University of Memphis, Memphis, TN and Central China Normal University, Wuhan, China

Abstract

The *BRCA Gist* Intelligent Tutoring System helps women understand and make decisions about genetic testing for breast cancer risk. *BRCA Gist* is guided by Fuzzy-Trace Theory, (FTT) and built using AutoTutor Lite. It responds differently to participants depending on what they say. Seven tutorial dialogues requiring explanation and argumentation are guided by three FTT concepts: forming gist explanations in one's own words, emphasizing decision-relevant information, and deliberating the consequences of decision alternatives. Participants were randomly assigned to *BRCA Gist*, a control, or impoverished *BRCA Gist* conditions removing gist explanation dialogues, argumentation dialogues, or FTT images. All *BRCA Gist* conditions performed significantly better than controls on knowledge, comprehension, and risk assessment. Significant differences in knowledge, comprehension, and fine-grained dialogue analyses demonstrate the efficacy of gist explanation dialogues. FTT images significantly increased knowledge. Providing more elements in arguments against testing correlated with increased knowledge and comprehension.

Keywords

Intelligent Tutoring System; Fuzzy-Trace Theory; Medical Decision-Making; Breast Cancer; Genetic Testing

1.1 Introduction

Shared decision-making among patients and health care providers has become the paradigm for medical decision-making. In matters of treatment, testing, and preventive care the ideal is for patients and providers to make decisions together (Col et al., 2011). Indeed, there have

Corresponding Author: Christopher R. Wolfe, Ph.D., Department of Psychology, Miami University, Oxford, OH 45056, WolfeCR@MiamiOH.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

been calls to more fully include patients in the process of making medical diagnoses (Graedon & Graedon, 2014). The professional expectation is that patients, physicians, and other professionals will collaboratively decide about the best course of action given the available medical evidence and the unique needs and values of each patient. Of course, everyone is potentially a medical patient, and patients rarely have medical training. Given the premium placed upon shared medical decision-making, there is an acute need for effective and efficient informal education for patients.

Breast cancer is one domain for which there is a significant need to help everyday women understand complex information and make informed decisions (Reyna, Nelson, Han, & Pignone, 2015). Patient education strategies for breast cancer include pamphlets and books, (e.g. Love, 2010) web sites, (e.g. National Cancer Institute, 2014) and patient testimonials and other narratives (Shaffer, Hulsey, & Zikmund-Fisher, 2013). Our approach has been to develop an Intelligent Tutoring System (ITS) called *BRCA Gist* (BREast CANcer and Genetics Intelligent Semantic Tutoring) to help healthy women understand and make decisions about genetic testing for breast cancer risk (Wolfe et al., 2015). There is solid evidence of the effectiveness of *BRCA Gist* (Wolfe et al., 2015; Wolfe et al., 2013; Widmer et al., 2015). The purpose of the current investigations is to isolate the processing loci responsible for effective learning, comprehension, and decision-making when women interact with this ITS.

Below we provide a brief overview of issues associated with decision-making about breast cancer and genetic risk, and describe the *BRCA Gist* ITS. We then present experimental data and detailed analyses of tutorial dialogues between women and *BRCA Gist* to help pinpoint the loci of the efficacy of the *BRCA Gist* system with respect to a theoretically-grounded form of self-explanation called gist explanation: the use of graphs and other specifically constructed images grounded in Fuzzy-Trace Theory, (FTT, Reyna, 2008a) and generating arguments for and against genetic testing for breast cancer risk.

1.2 Breast Cancer and Genetic Risk

Decisions about whether to be tested for genetic risk of breast cancer are difficult. Understanding risks and making good decisions requires health literacy and "numeracy" (Reyna & Brainerd, 2007; 2008) to interpret the meaning of base rates, joint probabilities (Wolfe & Reyna, 2010a,b) conditional probabilities, (Peters, McCaul, Stefanek, Nelson, 2006; Wolfe, Fisher & Reyna, 2012) and other quantitative concepts. Systematic biases in risk estimation have been demonstrated for both patients and providers (Offit, 2006; Reyna, Lloyd, & Whalen, 2001; Reyna, Nelson, Han, & Dieckmann, 2009). Women must also reason with ambiguous technical information in the context of conflicting attitudes and competing goals and constraints. BRCA testing accompanied by genetic counseling is expensive and without a family history of breast cancer, often not covered by insurance (Agus, 2013; Andrews, 2013). There are only about 3000 genetic counselors in the United States (Karow, 2013) to help women make these decisions.

Genetic testing for breast cancer risk potentially saves lives. However, because of the low base rate of BRCA mutations, the expense of testing which is often not covered by

insurance, and relatively high rate of ambiguous results, most women are not good candidates for predictive genetic testing for breast cancer risk. Unfortunately, little time is available for patients and physicians to discuss the complex issues surrounding genetic risk. Many patients are unsure what they would do if they received positive, negative, or ambiguous results. Yet those receiving positive results must decide about measures such as Tamoxifen treatments, more frequent mammograms, screening for ovarian cancer, and prophylactic mastectomy (Armstrong, Eisen, & Weber, 2000; Chao, Studts, Abell, Hadley, Roetzer, & Dineen et al., 2003; Stefanek, Hartmann, & Nelson, 2001) and negative results do not guarantee a lifetime free of cancer. Interest in genetic testing does not always coincide with assessed medical risk, and low-risk women are unlikely to consider all of the implications of testing. There are simply not enough genetic counselors to talk with every woman pondering genetic testing for BRCA mutations, highlighting the value of an effective and scalable ITS.

1.3.1 BRCA Gist

BRCA Gist engages women in a dialogue about many difficult issues associated with genetic testing for breast cancer risk (Armstrong, Eisen, & Weber, 2000; Berliner & Fay, 2007; Stefanek, Hartmann, & Nelson, 2001). Azevedo and Lajoie (1998) developed a prototype tutor to train radiology residents in diagnosing breast disease with mammograms. However, *BRCA Gist* appears to be the first use of any ITS in the domain of patients' medical decision-making. This approach is promising for helping laypeople understand and make decisions about breast cancer risk (Brewer, Richman, DeFrank, Reyna, & Carey, 2012) because individual one-on-one human tutoring is perhaps the best approach to facilitating deep conceptual understanding (Chi, Siler, Jeong, Yamauchi & Hausmann, 2005) with most human tutors yielding effect sizes of about .8, which is comparable to the best ITS (VanLehn, 2011) and truly expert tutors performing significantly better. Recent research on ITS has been very promising (du Boulay, 2016). In a recent meta-analysis of findings from 50 controlled evaluations of ITS, Kulik & Fletcher (2015) found that the median effect of intelligent tutoring in 50 evaluation studies was to increase test scores 0.66 standard deviations.

BRCA Gist is guided by Fuzzy-Trace Theory, (FTT) Reyna's (2008a) influential theory of medical decision-making grounded in basic research on memory and quantitative reasoning (e.g. Wolfe & Reyna, 2010b). From a FTT perspective, people are mainly gist processors, with the word "gist" used much as it is in everyday speech meaning the essential bottom-line meaning. FTT holds that, when information is encoded, people form multiple mental representations along a continuum from verbatim representations with superficial detail to fuzzier gist representations capturing the bottom-line meaning (Reyna, 2012; Reyna & Brainerd, 2011). Thus, gist and verbatim representations are formed in parallel during information acquisition. In decision-making, people prefer to reason with the vaguest bottom-line gist that can be used to decide among options (Reyna, Chick, Corbin, & Hsia, 2014; Wilhelms & Reyna, 2013). The preference to operate on the crudest gist, the fuzzy-processing preference, increases with experience or expertise (Reyna, 2008a; Reyna & Lloyd, 2006). In making decisions it is often more helpful to rely on these fuzzy gist representations (Reyna & Mills, 2014) provided they accurately capture decision-relevant

information. Superior medical decision makers appear to distill their experience into flexible gist representations, and gist representations are also associated with better decisions about risk and health among laypersons (Fraenkel et al., 2012; Mills, Reyna, & Estrada, 2008; Reyna, Estrada, DeMarinis, Myers, Stanis, & Mills, 2011; Reyna & Mills, 2014).

BRCA Gist allows us to apply several complementary ideas rooted in FTT to help people make good medical decisions. First, gist-based interventions (Reyna, 2008b) improve knowledge, understanding, and decision-making in medical contexts. Second, helping people explain the gist of complex medical information *in their own words* fosters learning and comprehension (Lloyd & Reyna, 2009). The overarching goal of our research has been to advance understanding of how women decide about predictive genetic testing and to develop the *BRCA Gist* ITS for women deciding about genetic testing and breast cancer risk.

BRCA Gist uses three female avatars (Moreno, Mayer, Spires, & Lester, 2001) of various apparent ethnicities with facial expressions and simulated facial movements, voice inflection, and conversational phrasing (Graesser, VanLehn, Rose, Jordan, & Harter, 2001). It converses with people, responding to what they type in a text box. It processes users' verbal input using Latent Semantic Analysis, (LSA) to provide appropriate feedback, *BRCA Gist* uses LSA to compare sentences entered by users to expectations texts (Graesser, Wiemer-Hastings, Wiemer-Hastings, Harter, 2000; Wolfe et al., 2012; Wolfe et al., 2013). *BRCA Gist* compares what people say to expectation texts that we developed using human verbal research data and refined through a series of development iterations (Wolfe et al., 2013). LSA permits *BRCA Gist* to assess this association and respond accordingly – even when participants explain the gist of key concepts in their own words, using different terms than those in the expectations texts.

BRCA Gist helps people to form useful gist representations (Reyna, 2008a) rather than drilling them on verbatim facts. This is accomplished by presenting concepts with explanations that highlight the essential meaning of information, as well as figures and videos conveying the bottom-line gist of core concepts, stripping away irrelevant detail. Gist representations of numerical concepts emphasize the gist of categorical risk (risky vs. not risky) and ordinal risk (lower vs. higher) (Zikmund-Fisher, 2013). *BRCA Gist* is made up of four modules on breast cancer and metastasis, risk factors, genetic mutation testing, and the consequences of testing. It provides didactic information interspersed with seven tutorial dialogues on topics including those requiring an explanation, for example "How do genes affect breast cancer risk?" and "What should someone do if she receives a positive result for genetic risk of breast cancer?" and those requiring argumentation such as "What is the case for (and against) genetic testing for breast cancer risk?" (Wolfe et al., 2015). The cognitive science literature provides good evidence that actively generating and elaborating on explanations of complex materials promotes understanding (VanLehn et al., 2007; Roscoe & Chi, 2008). After presenting didactic information *BRCA Gist* asks people questions and helps them form good gist explanations of key decision-relevant information, and arguments both for and against genetic testing. Gholson and colleagues (2008) found that learning is facilitated when materials are organized around questions that invite deep reasoning, even for vicarious learners, and Craig, Gholson, Brittingham, Williams, & Shubeck (2009) found

that explanations combined with questions are effective for low knowledge learners of Newtonian physics.

Figure 1 is a screen shot from a *BRCA Gist* tutorial dialogue. It shows an animated agent that has just asked the person the question orally and with screen text, "What is the case against genetic testing for breast cancer risk?" The person has responded by typing in the textbox, "Genetic testing is expensive and most people do not have BRCA mutations." As the person continues to add text she will receive verbal feedback from the avatar.

BRCA Gist was built using AutoTutor Lite, (Nye, Graesser, & Hu, 2014; Sullins, Craig, & Hu, 2015; Wolfe, Fisher, Reyna, & Hu, 2012; Wolfe et al., 2013) a web-based version of AutoTutor (Graesser, 2011; Graesser & McNamara, 2010) created by Xiangen Hu (Hu, Han, & Cai, 2008). *BRCA Gist* is platform independent and designed to handle large numbers of users simultaneously. It has a talking animated agent interface (Graesser & McNamara, 2010) and converses with users based on expectations using hints and elaboration prompts. AutoTutor Lite is the first web-based ITS platform that allows learners to interact with it through the use of natural language (English). The web-based AutoTutor Lite lacks some of the sophistication of dialogue scripts in stand alone ITS such as AutoTutor (see Graesser, 2011; Graesser, Chipman, Haynes, & Olney, 2005; Graesser & McNamara, 2010; Graesser, McNamara, & VanLehn, 2005; Kopp, Britt, Millis, & Graesser, 2012). However, AutoTutor Lite is more than adequate for implementing three key principles from FTT to help women understand and make decisions about genetic testing for breast cancer risk: first, the importance of helping women form gist explanations in their own words; second, a focus on decision-relevant dimensions of the knowledge domain; and third that at least some tutorial dialogues should focus on the risks and consequences of decision alternatives.

1.3.2 The Efficacy of *BRGA Gist*

Previous research subjected *BRCA Gist* to three multi-site randomized, controlled experiments with women at two universities, and field experiments with a community sample of women recruited in upstate New York and women recruited on-line (Wolfe et al., 2015; Wolfe, Reyna et al., 2013; Widmer et al., 2015). Participants were randomly assigned to *BRCA Gist*, the NCI web pages about breast cancer and genetic risk, or a control group receiving an irrelevant tutorial. This strategy controls for much of the same verbatim content presented on the NCI web site and for the process of engaging with a tutor.

Declarative Knowledge was assessed with a multiple-choice test described in section 2.2.2. Starting with experiment 2, *Gist Comprehension* (Wolfe et al., 2015) was assessed with an instrument measuring participants' understanding of the essential bottom line meaning—or gist—of knowledge (see section 2.2.3). Finally, participants received 12 *Risk Assessment Scenarios*, (Wolfe et al., 2015 a measure of applied risk-assessment accuracy; see section 2.2.4).

Table 1 presents key outcomes by experiment and condition. Participants in both experiments at both universities who were randomly assigned to the *BRCA Gist* condition scored significantly higher on percent correct declarative knowledge than the NCI group,

and both groups scored significantly higher than the control. In the field experiment, *BRCA Gist* and NCI groups again scored significantly higher than the control. Differences between *BRCA Gist* and NCI groups were significant for less highly educated participants without advanced degrees (i.e. a MA or Ph.D.). Effect sizes were large, $\eta^2 = 0.2332$. There were significant differences between sites, but the site by condition interaction was not significant: *BRCA Gist* participants scored significantly higher at all sites.

We found a comparable pattern of results for gist comprehension. In both laboratory and field experiments, *BRCA Gist* participants scored significantly higher than NCI participants and both scored significantly higher than controls. In both experiments, *BRCA Gist* participants more strongly endorsed agreement with true statements (and disagreement with false statements). These tasks do not require reasoning with specific verbatim facts, but do require thinking about the meaning of information. Effect sizes were large, $\eta^2 = 0.2694$.

For risk assessment accuracy, we consistently found that the *BRCA Gist* group scored significantly higher than the control group, and slightly but not significantly higher than the NCI group. Effect sizes were medium, $\eta^2 = 0.1359$. *BRCA Gist* was effective for all groups, but the NCI web site appears to be slightly more helpful for highly educated women whereas *BRCA Gist* appears more uniformly helpful across levels of education.

Having demonstrated the efficacy of *BRCA Gist* in several studies, the purpose of the current investigation was to isolate the loci of its effectiveness. Three theoretically-motivated aspects of *BRCA Gist* worthy of systematic research are (a) tutorial dialogues in which participants engage in gist self-explanation explanation (Lloyd & Reyna, 2009); (b) the use of graphs and other specifically constructed images grounded in Fuzzy-Trace Theory (FTT, Reyna, 2008a; Brust-Renck, Royer, & Reyna, 2013); and (c) generating arguments for and against genetic testing for breast cancer risk (Wolfe, Britt, & Butler, 2009).

1.4 Gist Explanation

The value of self explanation has been well documented in the literature on cognition and instruction for over two decades (Chi, 2000; Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Chi, Leeuw, Chiu, & LaVanher, 1994; VanLehn, Jones, & Chi, 1992). Research suggests that actively generating and elaborating explanations of material is more beneficial to learning than passively spending time with the material by reading or listening to lectures (Graesser, McNamara, & VanLehn, 2005). When learning in complex domains, particularly scientific and academic knowledge, self-explanations are thought to be pedagogically deep because people must learn to express causal and functional relationships rather than mechanically applying rote procedures (VanLehn, Jones, & Chi, 1992). However, when using ITS to promote decision-making, following FTT, we argue for the importance of facilitating gist explanations that organize the bottom-line meaning of decision-relevant causal and functional relationships (Lloyd & Reyna, 2009). Thus, in the case of breast cancer, it is helpful to understand how cancer grows and spreads, but even more important to have a gist understanding that cancer becomes deadly when it spreads and that catching cancer early before it spreads (whether through surrounding tissues, the lymphatic system or the circulatory system) greatly increases one's chances of survival. FTT also suggests that

medical decision-making can be improved through tutorial dialogues when people are asked to consider the consequences of decision alternatives. In the case of BRCA Gist, we ask people to explicitly consider in their own words what someone should do if she received a positive (and negative) test result for genetic breast cancer risk. Given the cost and other "down side risks" associated with testing, if a person cannot articulate what they might do differently in the event of a positive or negative test result then there is little reason for testing. It is also important for people to understand that a negative test result for BRCA 1 or 2 mutations do not appreciably reduce a woman's risk of breast cancer from the base rate.

1.5 Gist Provoking Images

FTT suggests that images can be used to help people form useful and appropriate gist representations (Brust-Renck, Royer, & Reyna, 2013). Figure 2 presents one example of a simple graph showing long-term survival rates for women diagnosed with breast cancer at different stages. Line graphs communicate gist-based representation of global patterns of magnitude (Brust-Renck, Royer, & Reyna, 2013). The precise verbatim percentages are not important, except when they imply qualitatively different outcomes. In making decisions about breast cancer prevention, including decisions about genetic testing, it is important to understand that when breast cancer is caught early most people survive and the earlier it is detected, the better one's odds of survival.

BRCA Gist presents Figure 3 to help people understand the bottom line meaning of relative risk and absolute risk. As the avatar explains, the square on the left indicates that 2 in 200 women (shown in red) are affected and the square on the right shows that 4 in 200 are affected (e.g., by genetic risk). The relative risk increases by 100% even though the absolute risk only increases from 1% to 2%. Concepts such as relative risk, absolute risk, and 5-year risk are confusing to many people, especially those low in numeracy (Reyna & Brainerd, 2007; 2008). Icon arrays and displaying icons in a systematic grouped fashion makes it easier to get the gist of relative magnitude (Brust-Renck, Royer, & Reyna, 2013). FTT suggests that images such as these along with 2x2 tables reduce processing interference from class inclusion by disentangling nested classes and at drawing attention to the appropriate denominators, which can help improve judgment and decision-making (Wolfe & Reyna, 2010b).

1.6 Argumentation

Argumentation has been used in patient education in relation to claims about breast cancer (Mackay, Schulz, Rubinelli, & Pithers, 2007; Rubinelli, Schulz, & Paolini, 2008). Research on verbal reasoning suggests that it is helpful to encourage people to closely attend to the connections between claims and supporting reasons (Britt, Kurby, Dandotkar, & Wolfe, 2008). Given the tendency to imprecisely represent specific argument predicates, it is easy for people to over generalize and make unwarranted assumptions. Research on the "myside" bias (Wolfe & Britt, 2008; Wolfe, Britt, & Butler, 2009), suggests that attention to both pro and con side arguments may help people avoid such pitfalls in decisions about genetic testing for breast cancer. Previous research suggests that a tutoring system can be used to facilitate skills in argumentation (Wolfe et al., 2009).

1.7 Hypotheses

The purpose of this research is to isolate the processing loci responsible for effective learning, comprehension, and risk assessment when women interact with *BRCA Gist*. The general approach was to randomly assign participants to the full *BRCA Gist*, a control group, or one of three impoverished versions of *BRCA Gist* removing tutorial dialogues in which people are assisted in making gist explanations, tutorial dialogues in which people generate arguments for and against genetic testing for breast cancer risk, and a version that removes nine graphics and an animated movie in which the avatar builds a 2×2 Table showing the relationship between BRCA mutations and breast cancer in the general population. In addition, we conducted detailed analyses of the tutorial dialogues to assess their quality, coverage of material, and the relationship between those dialogues and learning outcomes. The hypotheses, statistical tests, and the source of each prediction are presented in Table 2. Our first hypothesis is that all of the *BRCA Gist* groups will perform significantly better than the control group on declarative knowledge, gist comprehension, and categorical risk assessment due to the overall approach rooted in FTT. Our second hypothesis is that the no gist explanation group will perform significantly lower than the full *BRCA Gist* group on key measures of knowledge, gist comprehension, and medical risk assessment, providing evidence that part of the locus of success is the verbal interactions in making gist explanations. The third hypothesis is that participants assigned to the version of *BRCA Gist* without the nine FTT images and the animated 2×2 table video will perform significantly lower on measures of knowledge, gist comprehension, and medical risk assessment than those assigned to the full *BRCA Gist* group. This would provide evidence that viewing these images is partially responsible for the effectiveness of *BRCA Gist*. Hypothesis Four is that among participants who gave gist explanations, *BRCA Gist* will be judged to respond to participants appropriately and that the quality of those explanations measured by trained human judges will be positively correlated with their quality measured by the *BRCA Gist* semantic engine as CO (coverage of expectations) score. This would provide evidence that *BRCA Gist* is responding appropriately in tutorial dialogues. Our fifth hypothesis is that among participants who gave gist explanations, the greater the quality of those verbal gist explanations the better the learning outcomes on measures of knowledge, gist comprehension, and medical risk assessment. This would provide additional evidence that the locus of the effectiveness of *BRCA Gist* stems, in part, from the verbal interactions between participants and *BRCA Gist* in making gist explanations. A sixth hypothesis is that the no argumentation group will perform significantly lower than the *BRCA Gist* group on measures of knowledge, gist comprehension, and risk assessment with a corollary seventh hypothesis that among participants who made arguments, the greater the quality of those arguments in terms of covering materials and elements of argumentation (Wolfe, Brit & Butler, 2009) the better the learning outcomes. This would provide evidence that developing arguments for and against genetic testing for breast cancer risk would lead to better outcomes on declarative knowledge, gist comprehension, and risk assessment. Thus, we will be able to assess the loci of processes through both experimental manipulations and fine-grained analyses of tutorial dialogues.

2.1 Method

2.2.1 Participants

Participants were 252 undergraduate women recruited at a university in the Midwest and a university in the Northeast who received course credit for participating. We recruited only women because the risk of breast cancer is about 100 times greater for women than for men, and women are the target audience for *BRCA Gist*. Data from one participant was excluded because she did not complete the experiment. Recruitment criteria were that participants had to be women over the age of 18 who had not themselves had breast cancer. According to self-reports, the mean age of participants was 19.6 years (SD=6.6) with 21.8% Asian or Asian American, 8.7% Black or African American, 8.7% Latina, 58.7% White, 2.3% mixed ethnicity, 4% selecting "other ethnicity" and 3.6% not answering the question in non-mutually exclusive categories (i.e., Hispanic, Latina or Spanish was asked separately).

2.2.2 Experimental Conditions

Participants were randomly assigned to one of six experimental conditions. The first condition was the full *BRCA Gist* tutor as used in previous experiments (n=40). The second was another full version of *BRCA Gist* built with an improved version of AutoTutor Lite (n=40). This version permits more efficient transitions from one unit to the next and has improved authoring tools and other "back end" improvements. We also made some minor changes to the didactic tutorial, for example, improving the pronunciation of some words, fixing grammatical errors, and making minor changes in wording. The differences between these versions were small and non-significant. Thus, for the analyses reported below, both of these conditions are combined into a single *BRCA Gist* condition.

In the control condition (n=45), participants received a tutorial created using AutoTutor Lite about a topic not relevant to breast cancer, nutrition and exercise. The tutor is equally effortful and time consuming, but does not teach any of the materials about testing for genetic risk of breast cancer. The next three conditions systematically impoverish *BRCA Gist*. The No Gist Explanation condition (n=40) removes the five tutorial dialogues in which *BRCA Gist* helps people form explanations to the questions "what is breast cancer," "how does breast cancer grow and spread," "how do genes affect breast cancer risk," "what should someone do if she receives a positive result for genetic risk of breast cancer," and "what should someone do if she receives a negative result for genetic risk of breast cancer." References to these questions were also removed from the tutorial; otherwise, it was identical to the *BRCA Gist* condition. The No FTT Images condition (n=44) removed nine figures created following FTT and an animated video clip of the avatar talking the user through a 2 × 2 table on incidence of BRCA mutations and breast cancer and the relationship between the two in the general population (see Reyna & Brainerd, 2008; Wolfe & Reyna, 2010). References to these images were removed from the tutorial; otherwise, it was identical to the *BRCA Gist* condition. Figures 2 and 3 are examples of FTT images removed from the tutorial in this condition. Fifty other images and another brief video clip not designed following FTT principles remained as part of the tutorial. The No Argumentation condition (n=42) removed the two dialogues in which *BRCA Gist* helped people develop arguments for and against genetic testing in response to the questions, "what

is the case for genetic testing for breast cancer risk” and “what is the case against genetic testing for breast cancer risk.” References to these questions were also removed from the tutorial; otherwise, it was identical to the *BRCA Gist* condition.

2.2.3 Instruments¹

A medical expert vetted tutorial content and research instruments. Unanswered items were scored as incorrect (declarative knowledge, gist comprehension, and risk assessment described below).

2.2.3.1 Declarative Knowledge—We developed 52 four-alternative multiple-choice items on breast cancer, genetic risk, and genetic testing (Wolfe et al., 2015). Items were created corresponding to modules on breast cancer and how it spreads (16 items); quantitative concepts and genetic risk (15 items); mutations, genetic testing, and genetic risk (11 items); and consequences of genetic testing (10 items). To illustrate, three sample items are, "Breast Cancer usually forms in which parts of the breast? (answer: ducts and lobules)," "What is the goal of surveillance? (answer: to find cancer early when it is most treatable)," and "Which of the following is a risk factor for breast cancer? (answer: having larger areas of dense breast tissue on a mammogram; having your first menstrual period before age 12; and going through menopause after age 55)." Cronbach's alpha for the instrument is .88.

2.2.3.2 Gist Comprehension of Genetic Breast Cancer Risk (Wolfe et al., 2015)—We developed a 40 item 1–7 Likert-scale instrument measuring gist comprehension of important information about breast cancer and genetic testing. Gist comprehension items such as, "the greatest danger of dying from breast cancer is when it spreads to other parts of the body" express the gist of that knowledge – the essential bottom-line meaning. People can strongly endorse statements such as these without remembering the precise verbatim details. Interestingly, people can also recall the specific numbers without comprehending their bottom line meaning, a phenomenon known as verbatim-gist independence (Reyna & Brainerd, 2008). The item stem is stated at a general level such that verbatim information is not needed to answer the question. The response format permits degrees of agreement, with some items reversed scored. Cronbach's alpha for Gist Comprehension is 0.85.

2.2.3.3 Risk Assessment Scenarios (Wolfe et al., 2015)—Participants received 12 scenarios describing a woman with no risk factors or medium or high genetic breast cancer risk based on Pedigree Assessment Tool (PAT) scores of 0, 3–5, and 8–10 respectively. Each description includes a name, age, ethnicity, hometown, family health facts, and personal health facts. Scenarios were equated for age, range of words between 56 – 60; range of Flesch Reading Ease Scores between 56.9 – 62.9; and range of Flesch-Kincaid Grade Level Scores between 7.3 – 7.9. To illustrate, one high-risk scenario read: "Claire is an unattached 35 year-old New Yorker. She has a vegan diet and is an avid jogger. Her family is of Scottish-Irish heritage. Recently, her 51-year-old uncle Sean was diagnosed with cancer of the breast. Claire has several siblings and to the best of her knowledge, her uncle Sean is the

¹These instruments can be downloaded online from <http://mdm.sagepub.com/content/35/1/46/suppl/DC1>

only family member with breast cancer." Participants evaluated risk by categorizing degree of genetic breast cancer risk for each woman as low, medium, or high.

2.3.1 Tutorial Dialogues

Our purpose in analyzing the tutorial dialogues was to determine whether *BRCA Gist's* measurement of the similarity between participant answers and expectation texts is a reliable measure for the quality of those answers, whether the quality of answers predicts learning, and the extent to which *BRCA Gist* responded appropriately to participants' verbal input. We assessed the quality of the *BRCA Gist* tutorial dialogues' interactions with research participants using reliable scoring rubrics (Wolfe et al., 2013).

Our approach to assessing coverage of content and the accuracy of *BRCA Gist's* assessment of the quality of answers was to use the final CO score for the last sentence entered by each participant. This score represents *BRCA Gist's* assessment of the semantic similarity between the participant's answer and the expectation text. To see if the CO scores accurately measure the degree of content covered in an answer, we compared *BRCA Gist* final CO scores to scores obtained applying our rubrics blind to CO score. To determine whether rubric measures were reliable, two independent trained raters used the rubric to assess about one third of the answers. Thus, if the tutor is appropriately interpreting verbal inputs from the users, an answer given a high CO score as a measure of semantic similarity should contain more relevant content, as measured by the researcher applying a rubric, than in an answer given a lower CO score. Applying a conditional reliability procedure (Wolfe, Widmer et al., 2013) the two judges achieved .89 agreement with a range of .84 with responses to the question "what should someone do if she receives a negative result for genetic risk of breast cancer," to .94 for responses to the question "how does breast cancer grow and spread." To assess the effect of the dialogues on learning, we correlated coverage scores with scores on the 52-item multiple-choice test assessing declarative knowledge of genetic risk of breast cancer, gist comprehension scores, and risk assessment percent correct.

To assess the success of interactions between *BRCA Gist* and participants we judged the appropriateness of each response made by *BRCA Gist*. Responses were judged as either appropriate or inappropriate. We used a gist scoring procedure to make a judgment for each *BRCA Gist* response. Judgments were made only in relation to the user's previous statement and not the entirety of the dialogue. The appropriateness-of-responses criteria were that the tutor's response did each of the following: (a) encouraged elaboration, (b) flowed naturally from the previous input, and (c) responded correctly to the accuracy of the participant's input. To be rated as appropriate, the *BRCA Gist* response had to meet all three criteria (the criteria for inappropriate was failing to meet one or more of the appropriateness criteria). About one third of the responses were used to train the judges. Two raters independently made judgments about one third of the responses. In calculating reliability, we examined the proportion of responses the two judges agreed upon divided by the total number of responses and found .95 agreement.

An argument is, at minimum, a claim supported by a reason (Toulmin, 1958; Voss & Van Dyke, 2001; Wolfe, Britt & Butler, 2009) and an important question about the argumentation dialogues is whether participants actually make arguments. We subjected each of the

argumentation dialogues to an analysis using rubrics assessing the presence or absence of elements or argumentation (Cedillos-Whynott, Wolfe, Widmer, Brust-Renck, Weil, & Reyna, in press; Wolfe, Britt & Butler, 2009). Each argumentation dialogue was assigned a score from 0 to 4 where 0 = no reasons provided; 1 = reasons are listed or stated without any connection to a claim; 2 = claims and reasons are stated and the connection between them (the warrant) is implied but not stated; 3 = three or more argument elements – in addition to the warrant – are implied but not stated (examples include claim, reason, backing, counterargument, rebuttal); and 4 = three or more argument elements, in addition to the warrant, are explicitly stated. Thus, scores of 0 or 1 fail to meet the minimum definition of an argument.

Participants were recruited on-line and the experiment took place in the laboratory. In all conditions, interacting with the avatar took approximately 90 minutes. Participants took about 30 minutes to complete the dependent measures.

3.1 Results

For declarative knowledge (percent correct), there was a significant main effect for condition supporting Hypothesis 1, $F(4, 239) = 12.02, p < .0001$ (see Table 3), a significant main effect for location with participants at the Northeastern university scoring significantly higher (74.4%, $SD=14.9$) than the Midwestern university (63.1%, $SD=17.1$) $F(1, 239) = 35.8, p < .0001$; the location by condition interaction was not significant, $F < 1$. Planned comparisons between means were made using Hsu-Dunnnett Least-Squares Means tests reveal that the control group was significantly lower than all of the other groups (Hypothesis 1). The *BRCA Gist* group scored significantly higher on declarative knowledge than the No Gist-Explanation group (Hypothesis 2); $Q = 1.98, p = .037$. The *BRCA Gist* group also scored significantly higher than the No FTT Images group, $Q = 1.98, p = .049$; (Hypothesis 3, see Table 3).

For Gist Comprehension, there was a significant main effect for condition confirming Hypothesis 1, $F(4, 239) = 14.5, p < .0001$, (see Table 3) a significant main effect for location with participants at the Northeastern university scoring significantly higher (5.19, $SD=0.61$) than the Midwestern university (4.85, $SD=0.59$) $F(1, 239) = 23.6, p < .0001$; the location by condition interaction was not significant, $F(4, 239) = 1.07, p = .37$. Planned comparisons using the Hsu-Dunnnett Least-Squares Means test produced a borderline effect for Hypothesis 2 with the *BRCA Gist* group higher on gist comprehension than the No Gist-Explanation group, $Q = 1.98, p = .099$. Contrary to Hypothesis 3, No FTT Images was not significantly different than BRCA Gist, and the control group was significantly lower than all of the other groups (Hypothesis 1, see Table 3).

For the risk assessment scenarios, there was a significant main effect for experimental condition confirming Hypothesis 1, $F(4, 241) = 5.78, p < .0001$, there was an effect for location, $F(1, 241) = 16.06, p < .0001$, and the location by condition interaction was not significant, $F(4, 241) = 1.40, p = .23$ (see Table 3). Participants at the Northeastern university had a significantly higher percent correct, 60.5% ($SD=13.7$) than participants at the Midwestern university, 54.4% ($SD=13.1$). Planned comparisons using the Hsu-Dunnnett

Least-Squares Means test indicates that the *BRCA Gist* mean is significantly greater than the No Gist Explanation mean supporting Hypothesis 2, $Q=2.485$, $p = .01$, there was not a significant difference between the No FTT Images and *BRCA Gist* groups (Hypothesis 3), and the control group mean was significantly lower than the means for each of the other groups (see Table 3).

3.2 Tutorial Dialogues

Judges determined that *BRCA Gist* responded appropriately to verbal input from participants for 97.7% of tutorial responses supporting Hypothesis 4. This represents a substantial improvement over the 85% of responses judged to be appropriate in the first iteration of *BRCA Gist* reported by Wolfe, Widmer, and colleagues (2013) who also found that the percentage of appropriate responses correlated with learning outcomes. Perhaps because, in the current study, *BRCA Gist* performance approaches ceiling, the percentage of appropriate responses did not predict performance on the assessment instruments, $F_s < 1$.

In the three experimental conditions in which participants engaged in gist explanation dialogues, over the course of five gist explanations dialogues, participants produced a mean of 31.0 ($SD=10.8$) conversational turns (i.e. they typed an average of 31 sentences into the dialogue box when interacting with the avatar). Table 4 provides a breakdown by gist explanation question. In assessing the extent to which those gist explanations covered content, in the five gist explanation dialogues, there were a total of 75 items that were gist scored as present or absent by judges with the rubrics. Overall, participants covered a mean of 30.0% of the material outlined in the rubric ($SD=11.4$) with the breakdown by gist explanation question provided in Table 4. As indicated in Table 4, the rubric scores were highly correlated with the final internal CO coverage scores generated by *BRCA Gist* supporting Hypothesis 4 with correlations ranging from .63 to .84 and in each case $p < .0001$. This indicates that the *BRCA Gist* expectation texts for the gist explanation dialogues accurately assess the semantic content of participants' gist explanations with respect to tutorial content. Following Hypothesis 5, the number of rubric content items included in the gist explanation dialogues was a good predictor of declarative knowledge, $r = .368$, $p < .0001$ with greater the overall rubric scores associated with greater declarative knowledge scores. As shown in Table 4, this effect was significant for three of the dialogues on how genes affect breast cancer risk, and what to do in the event of positive and negative test results. This was also of borderline significance for the other two dialogues. Consistent with Hypothesis 4, content included in gist explanations was also a good predictor of gist comprehension scores, $r = .361$, $p < .0001$, with the greater the overall rubric score, the greater the gist comprehension score. As shown in Table 4, this effect held for each of the five gist explanation dialogues. However contrary to Hypothesis 4, gist explanation dialogues did not predict risk assessment percent correct, $r = .030$ $p = .74$, and was not significant for any of the five gist explanation dialogues (see Table 4).

Contrary to Hypothesis 6, the No Argument group was not appreciably different from the *BRCA Gist* group on any of the learning outcome variables (see Table 3). In making the case for testing, participants took a mean of 5.69 conversational turns ($SD=1.89$) and in making the case against genetic testing they took a mean of 6.69 conversational turns ($SD=2.70$). In

analyzing the verbal interactions asking people to make a case for genetic testing, we found significant correlations between rubric coverage scores and the final *BRCA Gist CO* coverage scores, $r = .319$, $p = .0051$. This is substantially lower than comparable correlations for the gist explanation dialogues. For the case against testing, the correlation between rubric and *BRCA Gist CO* score was more in line with the lowest gist explanation dialogue, $r = .637$, $p < .0001$. Over 31 content items on the pro and con argument content rubrics, participants covered a mean of 25.9% of the content ($SD=8.5$). For arguments in favor of testing, the mean percentage of content covered was 43.4% ($SD=18.3$) and for arguments against testing the mean of rubric coverage items was 19.8% ($SD=8.9$). Consistent with Hypothesis 7, the percentage of rubric coverage items included in the two argumentation dialogues predicted performance on the declarative knowledge test, $r = .243$, $p = .04$; and gist comprehension scores, $r = .247$, $p = .037$; and risk assessment percent correct was of borderline significance, $r = .206$, $p = .083$. However, neither the pro or con side argument alone produced rubric coverage scores that correlated with any of these learning outcomes at $p < .05$.

Following the procedure of Cedillos-Whynott and colleagues (in press) we subjected each argumentation dialogue to analysis with a rubric assessing whether each dialogue met the criteria for being an argument and the degree of sophistication in using elements of argumentation. Overall, only 46.5% of these dialogues met the minimum criteria for being an argument, operationalized as a claim supported by one or more reasons with the connecting warrant implied but not necessarily stated. This is comparable to the 43.7% found by Cedillos-Whynott and colleagues (in press). Overall, 1% earned a 0 meaning no reasons provided; 52.5% earned a 1 meaning reasons are stated or listed without any connection to a claim; 18.5% earned a 2 meaning claims and reasons are stated and the warrant is implied but not stated; 21% earned a 3 meaning three or more argument elements, (claim, reason, backing, counterargument, rebuttal etc. in addition to the warrant) are implied but not stated; and 7% earned a 4 meaning three or more argument elements (in addition to the warrant) are explicitly stated. Argumentation scores for arguments against, but not for, genetic testing were good predictors of learning outcomes. Consistent with Hypothesis 7, the greater the argumentation score in arguments against testing, the greater the declarative knowledge score, $r = .209$, $p = .035$; and the greater the gist comprehension score, $r = .265$, $p = .007$; but the correlation between argumentation score and risk assessment percent correct was not significant, $r = .155$, $p = .119$. Contrary to Hypothesis 7, the correlation between argumentation scores in dialogues in favor of testing and outcomes was not significant, for declarative knowledge, $r = -.019$, $p = .852$; for gist comprehension, $r = .040$, $p = .698$; and for risk assessment, $r = -.024$, $p = .817$.

In preparing *BRCA Gist* for gist explanation tutorial dialogues, we employed three principles rooted in FTT: that people should be encouraged to develop gist explanations in their own words rather than encouraging verbatim recitation of the presented text; that the dialogues should emphasize decision-relevant dimensions of genetic breast cancer risk; and that some dialogues should focus on the consequences of decision alternatives. The dialogue below exemplifies each of these ideas.

BRCA Gist: What should someone do if she tests positive?

Person: She should let her family members know and realize how much she is at risk for developing cancer if she tests positive for BRCA mutations.

BRCA Gist: Okay. Try writing about active surveillance.

Person: Active surveillance helps them find cancer at the earliest stage possible, if it is going to appear.

BRCA Gist: How about this. Try to add something about surgery.

Person: She could also have healthy breast tissue removed to decrease her chances of developing breast cancer.

BRCA Gist: Please write a little bit more.

Person: She could also have her ovaries or fallopian tubes removed in order to reduce her risk of ovarian cancer.

BRCA Gist: See if you can say something about drugs.

Person: There is also a drug she can take that can reduce her risk for cancer, even though it is not sure exactly how much it is preventative. These are all ways to be safe about breast cancer, finding this out.

BRCA Gist: Ok, let's move on. Please click on the finish button.

In this dialogue, the person has elected to start with notifying family members, and *BRCA Gist* prompts her to discuss active surveillance, surgery, and drugs and probes for more information by asking her to, "write a little bit more." The conversation focuses on consequences of testing in the event of a positive test result and the focus on what she should do encourages her to think about decision-relevant alternatives in her own words. Below are excerpts from six dialogues (i.e., six participants) in response to the question "what should someone do if she receives a negative result for genetic risk of breast cancer" exemplifying the focus on the consequences of decision alternatives (spelling mistakes and typos in the original).

BRCA Gist asks: Can you be a little more specific. What should someone do if she has a negative test result?

Person A: They should share this information with family members.

Person B: If they have a negative test result they should tell their family and children.

Person C: This does not mean that she is not at risk for breast cancer

BRCA Gist asks: You are doing very well. Can you say something about true negatives and false negatives?

Person D: True negatives mean that a test result comes back negative and the person does not actually have the genetic mutation, whereas a false negative means that the result comes back negative, but she actually has the mutation.

Person E: False negatives could happen due to testing errors, but true negatives would indicate that the person being tested has no higher of a risk than the general population.

Person F: A true negative means that a woman's risk of breast cancer is at the level of the general public.

Here, some participants emphasized that in the event of a negative test result the women would still be at risk of breast cancer, while others emphasized the consequences for other family members. Both of these are important considerations in decisions about genetic testing for breast cancer risk. These excerpts also illustrate that *BRCA Gist* can interact effectively even when confronted with spelling mistakes, poor grammar, and typos. The dialogue excerpts below further exemplify the concept of emphasizing decision-relevant dimensions for the question, "how do genes affect breast cancer risk?" (spelling mistakes and typos in the original).

BRCA Gist asks: ... can you say more about genetic risk factors?

Person G: Genetic risk factors include having a mother, sister, or other close relative with breast cancer, or being of Ashkenazi Jewish descent.

Person H: Risk factors increase if you have a family history of breast or ovarian cancer, especially if a male in the family has had breast cancer.

Person I: Just because someone has a risk of breast cancer it does not mean that will automatically get breast cancer, genetic risk factors serve as almost a heads up for a woman and her family members.

It is clear that different people emphasize different aspects of the tutorial as decision relevant. For example, it would not be surprising for a woman of Ashkenazi Jewish heritage to place greater emphasis on this dimension compared to a woman who is not Jewish, and a woman with (rare) male breast cancer in her family would be wise to give significant weight to this aspect of her family history.

4.1 Discussion

As in previous studies, all versions of *BRCA Gist* performed significantly better than the control group, providing evidence for the overall effectiveness of *BRCA Gist* including didactic text and other aspects of the tutorial (Hypothesis 1). However, of particular interest are the comparisons testing the effectiveness of the gist explanation dialogues, the images, and video clip designed following FTT, and the argumentation dialogues. We have strong evidence that the gist explanation dialogues improve learning outcomes (Hypothesis 2). The version of *BRCA Gist* with the 5 gist explanation dialogues yielded significantly higher scores on declarative knowledge, risk assessment, and there was a borderline significant effect for gist comprehension, compared to the version without these dialogues. Analyses of

the gist comprehension dialogues themselves provide evidence that *BRCA Gist* is accurately assessing coverage of content and responding appropriately, (Hypothesis 4). When learners include more content in their verbal responses they perform better on subsequent measures of knowledge and comprehension (Hypothesis 5). These results support the FTT predictions that helping people form useful gist representations improves risk assessment (Reyna, 2004; Reyna & Brainerd, 2008), comprehension (Reyna, Nelson, Han, & Dieckmann, 2009), and knowledge acquisition (Wolfe, Reyna, & Brainerd, 2005). "Drilling down," there is some further evidence for the effectiveness of helping people form gist explanations in their own words, discussing decision-relevant information, and considering the consequences of decision alternatives. Correlational evidence suggests that the most effective dialogues revolved around the decision-relevant question, "how do genes affect breast cancer risk" and the question about consequences, "what should someone do if she receives a positive result for genetic risk of breast cancer." Based on correlational evidence from the fine-grained analysis of gist explanation dialogues, the questions, "what is breast cancer," and "how does breast cancer grow and spread" were associated with smaller gains in outcome measures. These results seem to corroborate other findings on deep-level reasoning questions (Craig et al., 2012; Gholson et al., 2009). The effectiveness of asking participants to provide explanations in their own words is worthy of future experimental work.

There is also solid experimental evidence that the use of images rooted in FTT improved learning (Hypothesis 3). The nine FTT images and brief animated video clip of an animated 2×2 table was responsible for higher declarative knowledge scores. Given the amount of information in the 90 minute *BRCA Gist* tutorial including didactic text and tutorial dialogues, the improvements associated solely with graphs such as those shown in Figures 2 and 3 are theoretically telling. Each of these graphs was designed to help people form appropriate gist representations (Lloyd & Reyna, 2009; Wolfe et al., 2015). These results cannot be explained by the overall importance of graphics in ITS because fifty images and another brief video clip not designed following FTT principles were part of the tutorial in both the no FTT images and full *BRCA Gist* conditions.

There is no evidence that the tutorial dialogues asking people to make a case for, and a case against, genetic testing for breast cancer risk improved learning outcomes (Hypothesis 6). Replicating previous findings, (Cedillos-Whynott et al., in press) fine-grained analyses indicate that a majority of participants simply listed reasons – which fails to meet the minimum operational definition of an argument (Wolfe, Britt & Butler, 2009). However, when participants actually engaged in argumentation, they showed gains in knowledge and comprehension partially supporting Hypothesis 7. Both coverage of content in arguments and including more argumentation elements such as claims, reasons, backing counterargument, and rebuttal in arguments against testing yielded improvements in gist comprehension and declarative knowledge. An issue for future research is whether a different set of instructions and a different approach to the argumentation dialogues would help people form better arguments, resulting in better learning outcomes. It is possible but unlikely that the gist explanation dialogues were more effective than the argumentation dialogues simply because there were five explanation dialogues and only two argumentation dialogues. We found sharp differences between the pro and con side arguments as predictors of learning outcomes.

With respect to Hypothesis 7, a comparison of results from the current study to previous research on *BRCA Gist* conducted by Cedillos-Whynott and colleagues (in press) reveals some striking similarities. We found a comparable pattern for the percentage of rubric items covered in the two argumentation dialogues as predictors of outcomes. In the present study rubric coverage on the two argumentation questions predicted performance on the declarative knowledge test at $r = .243$ which was statistically significant but slightly lower than the $r = .323$ for pro side arguments and $r = .335$ for con side arguments found in previous research (Cedillos-Whynott, in press). For gist comprehension scores, we found $r = .247$ in the current study and $r = .311$ for pro side arguments and $r = .376$ for con side arguments found in previous research (Cedillos-Whynott in press). The percentage of rubric coverage items included in the two argumentation dialogues predicted performance on risk assessment percent correct at $r = .206$ which was of borderline significance compared to the $r = .236$ for pro side arguments and $r = .183$ for con side arguments found in previous research (Cedillos-Whynott, in press) that were both statistically significant. In the current research and earlier work we assessed the number of argument elements and assessed the relationship between argumentation scores and outcomes. In the present study, the greater the argumentation score in arguments against testing, the greater the declarative knowledge score, $r = .209$, which is less than the $r = .324$ found by Cedillos-Whynott et al. (in press) which was also statistically significant.

Taken collectively these results suggest that asking people to generate arguments is insufficient to produce gains in outcomes. Although facility with argumentation is a major goal of a university education (Wolfe, 2011) a large number of participants were unable to produce real arguments when asked to do so. Those who generated warranted arguments produced superior outcomes. However, it is apparent that more scaffolding teaching users how to develop an argument would be necessary to achieve the desired effects. Although this strategy may be useful in the context of learning in academic domains (Wolfe, Britt, Petrovic, Albrecht & Koop, 2009) in the context of helping patients make decisions about cancer risk there is insufficient evidence that argumentation is an effective strategy.

As in previous research, participants at the Northeastern university performed significantly better than those at the Midwestern university. This is not surprising because the former is more academically selective. Of greater importance, in the current study and previous research, we consistently find main effects without statistical interactions between location and experimental condition. This suggests that *BRCA Gist*, and constituent dimensions including gist explanation dialogues, FTT images, and argumentation dialogues are equally effective with different populations. In future research it will be important to include clinical samples of patients considering genetic testing for breast cancer risk.

There are a number of practical and theoretical questions that have not yet been addressed. One promising avenue is to explore the use of *BRCA Gist* as a tool for patient preparation by providing it to patients on a tablet in the waiting room before the clinical encounter with a physician or genetics counselor.

FTT suggests that having patients form gist explanations in their own words will be more effective than asking them to recall materials verbatim. However, differences between verbal

interactions among avatars and participants emphasizing gist or verbatim patient responses have not been explored systematically. There may also be more effective ways for *BRCA Gist* to encourage true argumentation including elements of argumentation such as counterarguments and rebuttal (Wolfe, et al., 2009).

BRCA Gist is the first ITS applied to patients' decision-making. It is unlikely to be the last. FTT provides a useful framework for understanding medical decision-making (Reyna, 2008a) and for developing effective decision tools (Fraenkel et al., 2012; Reyna, & Mills, 2014; Wolfe et al., 2015).

4.2 Conclusion

As predicted by FTT, *BRCA Gist* versions with five gist explanation dialogues yielded significantly higher scores on declarative knowledge, risk assessment, and gist comprehension, compared to the version without these dialogues, provides strong evidence that the gist explanation dialogues improve learning. When learners include more content in their verbal responses they perform better on subsequent measures of knowledge and comprehension. The use of nine images and a video clip of an animated 2×2 table rooted in FTT also improved learning. Evidence for the overall effectiveness of *BRCA Gist* stems from the finding that participants in all *BRCA Gist* conditions performed significantly better than the control group.

Acknowledgments

The project described was supported by Award Number R21CA149796 from the National Cancer Institute. 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 thank the National Cancer Institute for its support. We also wish to thank Rachel Aron, Andrew Circelli, Cecelia Favede, Jeremy Long, Mitch McDaniel, Ian Murphy, Kendall Powell, and Michael Thomas for capable assistance with data collection.

References

- Agus DB. The outrageous cost of a gene test. *New York Times*. [Accessed July 22, 2014] Available at http://www.nytimes.com/2013/05/21/opinion/the-outrageous-cost-of-a-gene-test.html?_r=0. Updated May 20, 2013.
- Andrews M. Coverage gaps can hamper access to some breast cancer screening, care. *Kaiser Health News*. [Accessed July 22, 2014] Available at <http://www.kaiserhealthnews.org/Features/Insuring-Your-Health/2013/052813-Michelle-Andrews-on-breast-cancer-care.aspx>. Updated May 27, 2013.
- Armstrong K, Eisen A, Weber B. Assessing the risk of breast cancer. *New England Journal of Medicine*. 2000; 342:564–571. [PubMed: 10684916]
- Azevedo R, Lajoie SP. The cognitive basis for the design of a mammography interpretation tutor. *International Journal of Artificial Intelligence in Education*. 1998; 9:32–44.
- Berliner JL, Fay AM, et al. Risk assessment and genetic counseling for hereditary breast and ovarian cancer: Recommendations of the National Society of Genetic Counselors. *Journal of Genetic Counseling*. 2007; 16:241–260. [PubMed: 17508274]
- Brewer NT, Richman AR, DeFrank JT, Reyna VF, Carey LA. Improving communication of breast cancer recurrence risk. *Breast Cancer Research and Treatment*. 2012; 133:553–561. [PubMed: 21964579]
- Britt MA, Kurby CA, Dandotkar S, Wolfe CR. I Agreed with What? Memory for Simple Argument Claims. *Discourse Processes*. 2008; 45:52–84.
- 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.

- du Boulay B. Recent Meta-reviews and Meta-analyses of AIED Systems. *International Journal of Artificial Intelligence in Education*. 2016; 26:536–537.
- Cedillos-Whynott EM, Wolfe CR, Widmer CL, Brust-Renck PG, Weil AM, Reyna VF. The Effectiveness of argumentation in tutorial dialogues with an intelligent tutoring system. *Behavior Research Methods*. (in press).
- Chao C, Studts JL, Abell T, Hadley T, Roetzer L, Dineen S, Lorenz D, YoussefAgha A, McMasters KM. Adjuvant chemotherapy for breast cancer: How presentation of recurrence risk influences decision-making. *Journal of Clinical Oncology*. 2003; 21:4299–4305. [PubMed: 14581440]
- Chi MT. Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. *Advances in instructional psychology*. 2000; 5:161–238.
- Chi MTH, Bassok M, Lewis MW, Reimann P, Glaser R. Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*. 1989; 15:145–182.
- Chi MTH, de Leeuw N, Chiu MH, LaVancher C. Eliciting Self-Explanations Improves Understanding. *Cognitive Science*. 1994; 18:439–477.
- Chi MTH, Siler SA, Jeong H, Yamauchi T, Hausmann RG. Learning from human tutoring. *Cognitive Science*. 2005; 25:471–533.
- Col N, Bozzuto L, Kirkegaard P, Koelewijn-van Loon M, Majeed H, Ng CJ, Pacheco-Huergo V. Interprofessional education about shared decision-making for patients in primary care settings. *Journal of Interprofessional Care*. 2011; 25:409–415. ISSN: 1356-1820. [PubMed: 22011026]
- Craig SD, Gholson B, Brittingham JK, Williams JL, Shubeck KT. Promoting vicarious learning of physics using deep questions with explanations. *Computers & Education*. 2012; 58:1042–1048.
- Fraenkel L, Peters E, Charpentier P, Olsen B, Errante L, Schoen R, Reyna VF. A decision tool to improve the quality of care in Rheumatoid Arthritis. *Arthritis Care & Research*. 2012; 64(7):977–985. [PubMed: 22392766]
- Gholson B, Witherspoon A, Morgan B, Brittingham JK, Coles R, Graesser AC, Sullins J, Craig SD. Exploring the deep-level reasoning questions effect during vicarious learning among eighth to eleventh graders in the domains of computer literacy and Newtonian physics. *Instructional Science*. 2009; 37:487–493.
- Graedon T, Graedon J. Let patients help with diagnosis. *Diagnosis*. 2014; 1(1):49–51.
- Graesser AC. Learning, thinking, and emoting with discourse technologies. *American Psychologist*. 2011; 66(8):746–757. [PubMed: 22082403]
- Graesser AC, Chipman P, Haynes BC, Olney A. AutoTutor: An intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions on Education*. 2005; 48(4):612–618.
- Graesser A, McNamara D. Self-regulated learning in learning environments with pedagogical agents that interact in natural language. *Educational Psychologist*. 2010; 45:234–244.
- Graesser AC, McNamara DS, VanLehn K. Scaffolding deep comprehension strategies through Point&Query, AutoTutor, and iSTART. *Educational Psychologist*. 2005; 40(4):225–234.
- Graesser AC, Wiemer-Hastings P, Wiemer-Hastings K, Harter D. Tutoring Research Group, & Person, N. Using latent semantic analysis to evaluate the contributions of students in AutoTutor. *Interactive Learning Environments*. 2000; 8(2):129–147.
- Hu, X.; Han, L.; Cai, Z. Semantic decomposition of student’s contributions: an implementation of LCC in AutoTutor Lite. Paper presented to the Society for Computers in Psychology; November 13, 2008; Chicago, Illinois. 2008.
- Karow J. As genomics increases the complexity of diagnostic tests, the role of genetic counselors expands. *Clinical Sequencing News*. [Accessed July 22, 2014] Available at <http://www.genomeweb.com/sequencing/genomics-increases-complexity-diagnostic-tests-role-genetic-counselors-expands>. Updated December 18, 2013.
- Kopp KJ, Britt MA, Millis K, Graesser AC. Improving the efficiency of dialogue in tutoring. *Learning and Instruction*. 2012; 22(5):320–330.
- Kulik JA, Fletcher JD. Effectiveness of Intelligent Tutoring Systems A Meta-Analytic Review. *Review of Educational Research*. 2015; 86:42–78.
- Lloyd FJ, Reyna VF. Clinical gist and medical education: connecting the dots. *Journal of American Medical Association*. 2009; 302:1332–1333.

- Love, SM. Dr. Susan Love's Breast Book. Fifth. Cambridge, MA: Da Capo Press; 2011.
- Mackay J, Schulz P, Rubinelli S, Pithers A. Online patient education and risk assessment: Project OPERA from cancerbackup: Putting inherited breast cancer risk information into context using argumentation theory. *Patient Education and Counseling*. 2007; 67:261–266. [PubMed: 17590305]
- Mills B, Reyna VF, Estrada S. Explaining contradictory relations between risk perception and risk taking. *Psychological Science*. 2008; 19:429–433. [PubMed: 18466401]
- Breast Cancer Risk in American Women. [Accessed July 23, 2014] National Cancer Institute. Available at <http://www.cancer.gov/cancertopics/factsheet/detection/probability-breast-cancer>. Updated September 24, 2012
- Moreno R, Mayer RE, Spires HA, Lester JC. The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction*. 2001; 19:177–213.
- Nye BD, Graesser AC, Hu X. AutoTutor and Family: A Review of 17 Years of Natural Language Tutoring. *International Journal of Artificial Intelligence in Education*. 2014; 24:427–469.
- Offit K. BRCA mutation frequency and penetrance: new data, old debate. *Journal of National Cancer Institute*. 2006; 98:1675–1677.
- Peters E, McCaul KD, Stefanek M, Nelson WA. Heuristics approach to understanding cancer risk perception: Contributions from judgment and decisionmaking research. *Annals of Behavioral Medicine*. 2006; 31:45–52. [PubMed: 16472038]
- Reyna VF. How people make decisions that involve risk. A dual-processes. approach. *Current Directions in Psychological Science*. 2004; 13:60–66.
- Reyna VF. A theory of medical decision-making and health: Fuzzy trace theory. *Medical Decision-making*. 2008a; 28:850–865. [PubMed: 19015287]
- Reyna VF. Theories of medical decision-making and health: An evidence-based approach. *Medical Decision-making*. 2008b; 28:829–833. [PubMed: 19020341]
- Reyna VF. A new intuitionism: Meaning, memory, and development in fuzzy-trace theory. *Judgment and Decision-making*. 2012; 7:332–359. [PubMed: 25530822]
- Reyna VF, Brainerd CJ. The importance of mathematics in health and human judgment: numeracy, risk communication, and medical decision-making. *Learning and Individual Differences*. 2007; 17:147–159.
- Reyna VF, Brainerd CJ. Numeracy, ratio bias, and denominator neglect in judgments of risk and probability. *Learning and Individual Differences*. 2008; 18:89–107.
- Reyna VF, Brainerd CJ. Dual processes in decision-making and developmental neuroscience: A fuzzy-trace model. *Developmental Review*. 2011; 31:180–206. [PubMed: 22096268]
- Reyna VF, Chick CF, Corbin JC, Hsia AN. Developmental reversals in risky decision-making: Intelligence agents show larger decision biases than college students. *Psychological Science*. 2014; 25:76–84. [PubMed: 24171931]
- 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; 37:1125–1142.
- Reyna VF, Lloyd F, Whalen P. Genetic testing and medical decision-making. *Archives of Internal Medicine*. 2001; 161:2406–2408. [PubMed: 11700152]
- Reyna VF, Mills BA. Theoretically motivated interventions for reducing sexual risk taking in adolescence: A randomized controlled experiment using fuzzy-trace theory. *Journal of Experimental Psychology: General*. 2014 Advance online publication. <http://dx.doi.org/10.1037/a0036717>.
- Reyna VF, Nelson WL, Han PK, Dieckmann NF. How numeracy influences risk comprehension and medical decision-making. *Psychological Bulletin*. 2009; 135:943–973. [PubMed: 19883143]
- Reyna VF, Nelson WL, Han PK, Pignone MP. Decision-making and cancer. *American Psychologist*. 2015; 9:122–127.
- Roscoe RD, Chi MTH. Tutor learning: The role of explaining and responding to questions. *Instructional Science*. 2008; 36:321–350.

- Rubinelli S, Schulz PJ, Paolini P. Argumentation in good news communication on genetic breast cancer. The experience of OPERA. Proceedings CMNA. 2008 [Accessed July 24, 2014] <http://www.cmna.info/CMNA8/programme/CMNA8-Rubinelli-etal.pdf>. Updated 2008.
- Shaffer VA, Hulsey L, Zikmund-Fisher BJ. The effects of process-focused versus experience-focused narratives in a breast cancer treatment decision task. *Patient Education & Counseling*. 2013; 93:255–264. [PubMed: 23962673]
- Stefanek M, Hartmann L, Nelson W. Risk-reduction mastectomy: Clinical issues and research needs. *Journal of the National Cancer Institute*. 2001; 93:1297–1306. [PubMed: 11535704]
- Sullins J, Craig SD, Hu X. Exploring the effectiveness of a novel feedback mechanism within an intelligent tutoring system. *International Journal of Learning Technology*. 2015; 10:220–236.
- Toulmin, S. *The uses of argument*. New York: Cambridge University Press; 1958.
- VanLehn K. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychology*. 2007; 46:197–221.
- VanLehn K, Graesser AC, Jackson GT, Jordan P, Olney A, Rose CP. When are tutorial dialogues more effective than reading? *Cognitive Science*. 2007; 31:3–62. [PubMed: 21635287]
- VanLehn K, Jones RM, Chi, & MTH. A model of the self-explanation effect. *Journal of the Learning Sciences*. 1992; 2:1–59.
- Voss JF, Van Dyke JA. Argumentation in psychology: Background comments. *Discourse Processes*. 2001; 32:89–111.
- 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. [PubMed: 25921818]
- Wilhelms EA, Reyna VF. Fuzzy trace theory and medical decisions by minors: Differences in reasoning between adolescents and adults. *Journal of Medicine and Philosophy*. 2013; 38:268–282.
- Wolfe CR. Argumentation across the curriculum. *Written Communication*. 2011; 28:193–219.
- Wolfe CR, Britt MA. Locus of the my-side bias in written argumentation. *Thinking & Reasoning*. 2008; 14:1–27.
- Wolfe CR, Britt MA, Butler JA. Argumentation schema and the myside bias in written argumentation. *Written Communication*. 2009; 26:183–209.
- 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. [PubMed: 19587180]
- Wolfe CR, Fisher CR, Reyna VF. Semantic coherence and inconsistency in estimating conditional probabilities. *Journal of Behavioral Decision-making*. 2012
- 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; 42:366–372. [PubMed: 20479168]
- Wolfe CR, Reyna VF. Semantic coherence and fallacies in estimating joint probabilities. *Journal of Behavioral Decision-making*. 2010b; 23:203–223.
- Wolfe, CR.; Reyna, VF.; Brainerd, CJ. *Transfer of Learning from a Modern Multidisciplinary Perspective* (p. 53–88). Greenwich, CT: Information Age Press; 2005. *Fuzzy-Trace Theory: Implications for Transfer in Teaching and Learning*.
- Wolfe, CR.; Reyna, VF.; Brust-Renck, PG.; Weil, AM.; Widmer, CL.; Cedillos, EM.; Fisher, CR.; Damas Vannucchi, I.; Circelli, AM. Efficacy of the *BRCA Gist* Intelligent Tutoring System to Help Women Decide About Testing for Genetic Breast Cancer Risk. Paper presented to the 35th Annual Meeting of the Society for Medical Decision-making; Baltimore, MD. 2013 Oct..
- Wolfe, CR.; Reyna, VF.; Widmer, CL.; Cedillos, EM.; Fisher, CR.; Brust-Renck, PG.; Chaudhry, S.; Damas Vannucchi, I. Efficacy of a Web-Based Intelligent Tutoring System on Genetic Testing For Breast Cancer Risk. Presentation to the 6th Annual Scientific Meeting of the International Society for Research on Internet Interventions; Chicago, IL. 2013.
- 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; 35:46–59. doi 0272989X14535983. [PubMed: 24829276]

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. [PubMed: 23709166]

Zikmund-Fisher BJ. The right tool is what they need, not what we have a taxonomy of appropriate levels of precision in patient risk communication. *Medical Care Research and Review*. 2013; 70:37S–49S. [PubMed: 22955699]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Highlights for Review

- Experimental evidence for the efficacy of brief gist explanation dialogues in an ITS.
- Experimental evidence for the efficacy of theoretically-grounded images in an ITS.
- Evidence that the judgments of the quality of user explanations made by an ITS and trained human judges are highly correlated.
- Evidence for the loci of the effectiveness of the first ITS for patients medical decision making.

What is the case against genetic testing for breast cancer risk?

Type Your Answer Below:

Genetic testing is expensive and most people do not have BRCA mutations.

O... Cu... Re... Ir... Ir... R...

Total Coverage

Attempts	Score
1	1.0

SCORE

Attempts

2/2

Finish Edit This SKO Create and Share Author's Email

Figure 1. Screen shot of *BRCA Gist* using the text box to answer the question, "What is the case against genetic testing for breast cancer risk?"

Percent Who Survive Each Stage of Breast Cancer

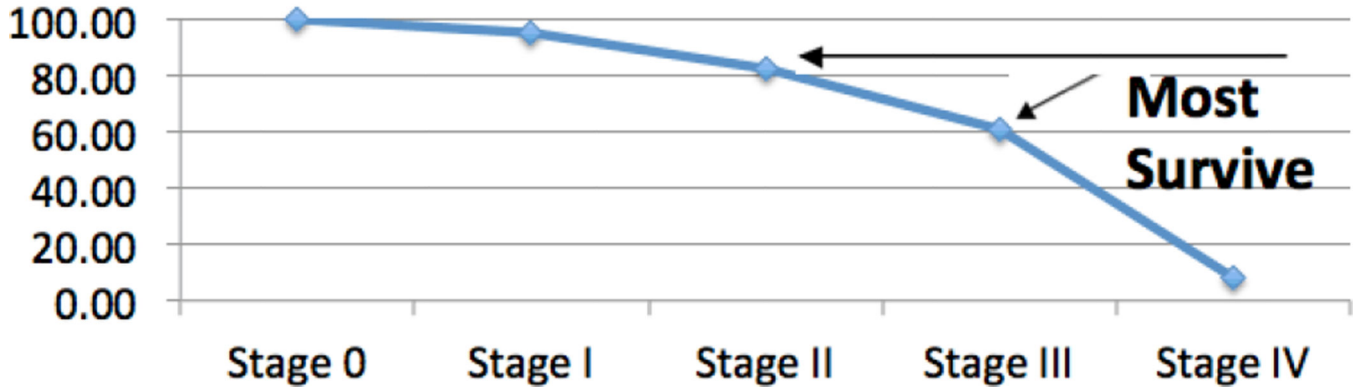


Figure 2.
Image capturing the gist that "most survive" breast cancer stages 0–3.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

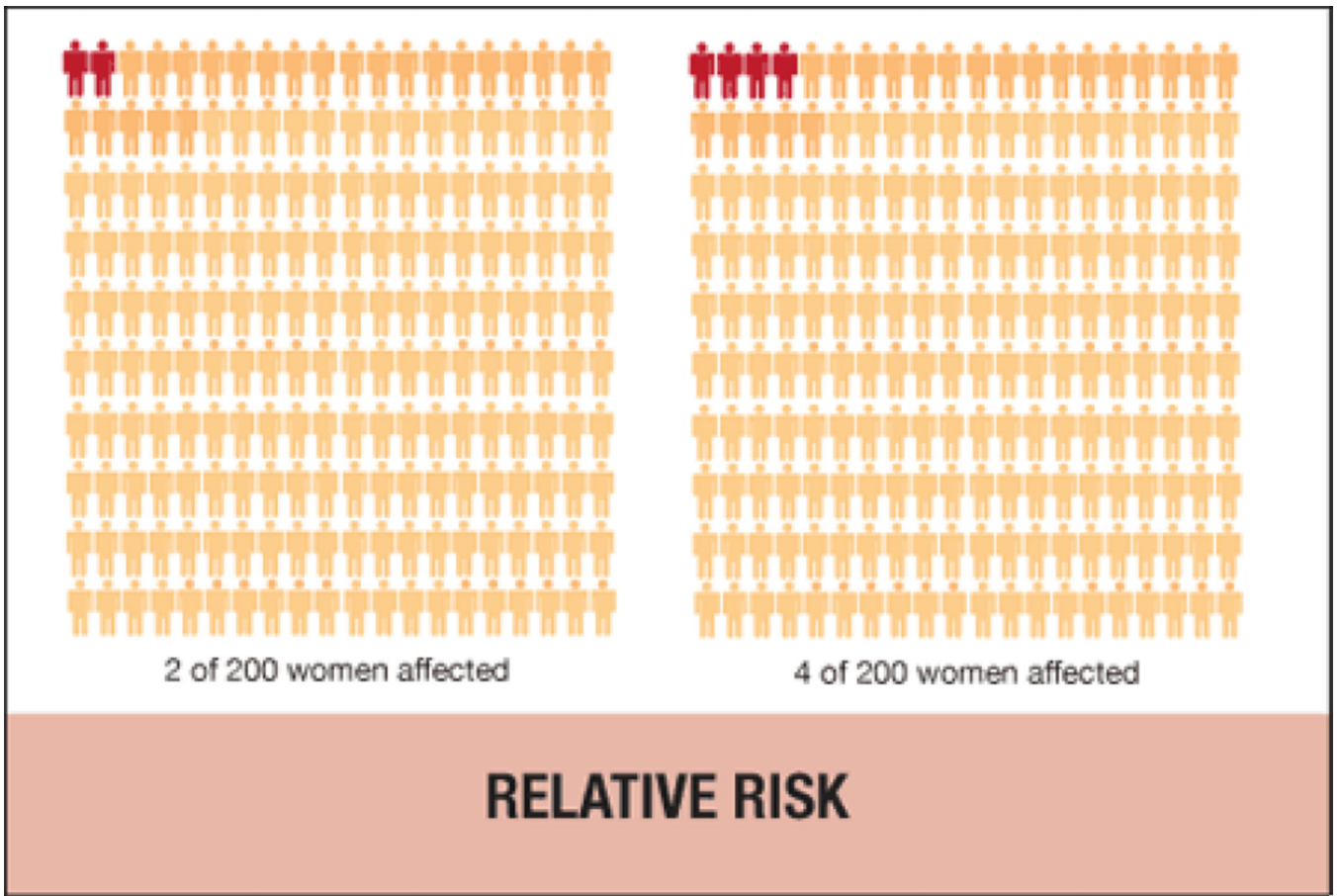


Figure 3. Graphic based on Fuzzy-Trace Theory to help people develop a gist understanding of the difference between relative and absolute risk.

Table 1

Means and standard deviations for key outcomes by condition and experiment (Wolfe, Reyna, Brust-Renck et al., 2013; Wolfe, Reyna, Cedillos et al., 2013; Wolfe et al., 2014).

Study	Experimental Condition	Knowledge Percent Correct	Gist Comprehension	Risk Assessment Percent Correct
University Lab Experiment 1	BRCA Gist	74% [*] (16)	-	59.6% [†] (16.5)
	NCI	67% (14)	-	55.4% (15.2)
	Control	56% (13)	-	46.8% (12.7)
University Lab Experiment 2	BRCA Gist	75% [*] (17)	5.34 [*] (0.68)	61.3% [†] (15.7)
	NCI	67% (14)	4.98 (0.42)	56.8% (15.7)
	Control	55 (15)	4.51 (0.57)	47.6% (11.7)
Web & Community Field Experiment	BRCA Gist	77% ^{†***} (17)	5.63 [*] (.69)	59% [†] (14.3)
	NCI	67% (20)	5.21 (.81)	56% (16.1)
	Control	57% (25)	4.60 (.55)	49% (14.3)

^{*} *BRCA Gist* > NCI Web and Control, $p < 0.0001$.

[†] *BRCA Gist* > Control, $p < 0.001$;

^{**} Excluding advanced degrees, *BRCA Gist* > NCI Web, Control, $p < 0.0001$.

Table 2

Hypotheses with statistical test and supporting principles of Fuzzy-Trace Theory.

	Hypotheses	Statistical Test	Source of Prediction
1	All four <i>BRCA Gist</i> versions will produce superior performance on declarative knowledge, gist comprehension, and categorical risk assessment, compared to the control group.	F and Hsu-Dunnett to Control Group	Replication and Extension
2	The no-gist-explanation group will perform significantly lower than the full <i>BRCA Gist</i> group on declarative knowledge, gist comprehension, and categorical risk assessment.	F and Hsu-Dunnett to Full <i>BRCA Gist</i> Group	FTT: Efficacy of Gist Self Explanation
3	The no-FTT-images group will perform significantly lower than the full <i>BRCA Gist</i> group on declarative knowledge, gist comprehension, and categorical risk assessment.	F and Hsu-Dunnett to Full <i>BRCA Gist</i> Group	FTT: Gist Representation of Images
4	In gist explanation tutorial dialogues, <i>BRCA Gist</i> will respond appropriately and participant coverage of expectations (measured by <i>BRCA Gist</i> and trained human judges using rubrics) will be positively correlated.	Correlation R between CO scores and rubric scores	Replication of Accuracy of <i>BRCA Gist</i> Semantic Engine
5	In gist explanation tutorial dialogues, participant coverage of expectations will be positively correlated with declarative knowledge, gist comprehension, and categorical risk assessment.	Correlation R between CO scores and outcome measures	FTT: Efficacy of Gist Self Explanation
6	The no argumentation group will perform significantly lower than the full <i>BRCA Gist</i> group on declarative knowledge, gist comprehension, and categorical risk assessment.	F and Hsu-Dunnett to Full <i>BRCA Gist</i> Group	Argument Schema Theory
7	In argumentation tutorial dialogues, participant argumentation scores and coverage of expectations will be positively correlated with declarative knowledge, gist comprehension, and categorical risk assessment.	Correlation R between CO scores, argument scores, and outcome measures	Argument Schema Theory

Table 3

Learning outcomes by experimental condition: mean outcome score with standard deviation in parentheses.

	Declarative Knowledge	Gist Comprehension	Risk Assessment Percent Correct
<i>BRCA Gist</i>	76.6% (12.6)	5.22 (0.58)	62.7% (14.1)
No Argumentation	76.7% (11.5)	5.22 (0.57)	59.7% (13.1)
No FTT Images	73.6% * (15.2)	5.15 (0.58)	59.9% (12.5)
No Gist Explanations	69.8% * (17.4)	4.98 (0.59)	56.5% * (12.4)
Control	56.1% † (15.3)	4.49 † (0.44)	49.4% † (12.1)

* Significantly lower than the *BRCA Gist* condition, $p < .05$.

† Control is significantly lower than all other conditions $p < .05$.

Table 4
Gist explanation dialogue coverage of content by question and correlations with learning outcomes.

Dialogue Question	Mean Participant Dialogue Conversation Turns i.e. sentences (SD)	Mean Percent Coverage of Rubric Content (SD)	Correlation BKCA Gist CO Score and Rubric Coverage Score	Correlation Rubric Coverage Score and Declarative Knowledge Percent Correct	Correlation Rubric Coverage Score and Gist Comprehension (p value)	Correlation Rubric Coverage Score and Risk Assessment Percent Correct
What is breast cancer?	5.9 (SD=2.7)	42.0% (SD=22.3)	r = .844 p<.0001	r = .149 p=.097	r = .240 p=.007	r = -.039 p=.663
How does breast cancer grow and spread?	5.7 (SD=2.5)	16.3% (SD=8.0)	r = .627 p<.0001	r = .160 p=.074	r = .227 p=.011	r = .061 p=.497
How do genes affect breast cancer risk?	7.1 (2.8)	20.0% (SD=10.3)	r = .699 p<.0001	r = .299 p=.0007	r = .237 p=.007	r = .100 p=.264
What should someone do if she receives a positive result for genetic risk of breast cancer?	6.0 (SD=3.2)	22.4% (SD=19.2)	r = .764 p<.0001	r = .289 p=.0004	r = .297 p=.0003	r = -.026 p=.758
What should someone do ... negative result for genetic risk of breast cancer?	6.2 (3.1)	30.8% (SD=21.4)	r = .826 p<.0001	r = .253 p=.002	r = .204 p=.014	r = .104 p=.210