Published in final edited form as:
J Exp Psychol Gen. 2014 August ; 143(4): 1627–1648. doi:10.1037/a0036717.

Theoretically Motivated Interventions for Reducing Sexual Risk Taking in Adolescence: A Randomized Controlled Experiment Applying Fuzzy-trace Theory

Valerie F. Reyna and
Cornell University

Britain A. Mills
University of Texas School of Public Health, Dallas

Abstract

Fuzzy-trace theory is a theory of memory, judgment, and decision-making, and their development. We applied advances in this theory to increase the efficacy and durability of a multicomponent intervention to promote risk reduction and avoidance of premature pregnancy and STIs. 734 adolescents from high schools and youth programs in three states (Arizona, Texas, and New York) were randomly assigned to one of three curriculum groups: RTR (*Reducing the Risk*), RTR+ (a modified version of RTR using fuzzy-trace theory), and a control group.

We report effects of curriculum on self-reported behaviors and behavioral intentions plus psychosocial mediators of those effects, namely, attitudes and norms, motives to have sex or get pregnant, self-efficacy and behavioral control, and gist/verbatim constructs.

Among 26 outcomes, 19 showed an effect of at least one curriculum relative to the control group: RTR+ produced improvements for 17 outcomes and RTR produced improvements for 12 outcomes. For RTR+, two differences (for perceived parental norms and global benefit perception) were confined to age, gender, or racial/ethnic subgroups. Effects of RTR+ on sexual initiation emerged six months after the intervention, when many adolescents became sexually active. Effects of RTR+ were greater than RTR for nine outcomes, and remained significantly greater than controls at one-year follow-up for 12 outcomes. Consistent with fuzzy-trace theory, results suggest that, by emphasizing gist representations, which are preserved over long time periods and are key memories used in decision-making, the enhanced intervention produced larger and more sustained effects on behavioral outcomes and psychosocial mediators of adolescent risk-taking.

Keywords

risk taking; risky decision making; adolescent health; STI prevention

Correspondence concerning this article should be addressed to Valerie F. Reyna, Human Neuroscience Institute, G331, Martha Van Rensselaer Hall, Cornell University, Ithaca, NY 14853. vr53@cornell.edu.
Valerie F. Reyna, Departments of Human Development and Psychology, Human Neuroscience Institute, and Center for Behavioral Economics and Decision Research, Cornell University; Britain A. Mills, University of Texas School of Public Health, Dallas

Sexually transmitted infections (STIs) and premature pregnancy are significant public health problems in the United States, and youth are at disproportionate risk (Centers for Disease Control and Prevention, 2012). For example, among the over 19 million new cases of STIs each year, almost half affect youth aged 15–24 (e.g., Kirby & Laris, 2009; Weinstock, Berman & Cates, 2004). In addition, despite declines, adolescent birthrates in the United States remain high, about six times the rates for adolescents age 15–19 in Japan or The Netherlands (e.g., National Campaign to Prevent Teen and Unplanned Pregnancy, 2013; Satterwhite et al., 2013). The negative consequences for individuals and society of early sexual initiation have been well documented, and extend to the children of mothers younger than 18 (e.g., Hoffman, 2006; Jemmott, Jemmott, & Fong, 2010; Martin et al., 2012).

These disturbing statistics are the result of risky behaviors, which are more prevalent in youth (e.g., IOM & NRC, 2011; Reyna & Farley, 2006). For example, adolescents have more sexual partners and are more likely to engage in unprotected sex; when they use protection, they are more likely to do so incorrectly or inconsistently, compared to adults (e.g., Abma et al., 1997; Satterwhite et al., 2013; Suellentrop, 2006). Hence, curriculum-based sex and STI education programs have been developed either to delay sexual initiation or, if youth are sexually active, to reduce risky behaviors (e.g., to reduce the number of sexual partners). Although programs initially focused on providing information, other psychosocial components have been added, such as values clarification and self-efficacy (e.g., improving communication skills used to refuse sex through role playing; Kirby, Barth, Leland, & Fetro, 1991; Reyna, Adam, Poirier, LeCroy, & Brainerd, 2005).

Currently, most programs rely on some variation of the expectancy-value framework, including the theory of reasoned action, prototype-willingness model, and behavioral decision-making approach (e.g., Fischhoff, 2008; Fishbein, 2008; Gerrard, Gibbons, Houlihan, Stock, & Pomery, 2008; Kirby, Coyle, Alton, Rolleri, & Robin, 2011). The essential elements of this framework are knowledge, perceptions of risk, attitudes, skills (e.g., self-efficacy to refuse sex), perceptions of norms (values), and intentions (Kirby et al., 2011). Although elements of this framework have been elaborated in new ways, theories of prevention and behavior change have not changed fundamentally for decades (National Institutes of Health, 2009).

In this article, we report the results of a new approach to reducing risky sexual behaviors in youth, a curriculum based on fuzzy-trace theory (e.g., Reyna, 2008). Fuzzy-trace theory is a theory of memory, decision making, and their development that is supported by the results of many experiments and mathematical models (e.g., Johns, Jones & Mewhort, 2012; Reyna, 2012a). The central concept of the theory is the distinction between two kinds of mental representations of information: gist (bottom-line meaning) and verbatim (literal facts). According to the theory, most adults base their judgments and decisions on the gist of information, rather than on the verbatim facts (for reviews, see Reyna, 2008; Reyna & Brainerd, 1995). In contrast, adolescents, especially those who take risks, are predicted to rely more on analysis of verbatim facts (Mills, Reyna, & Estrada, 2008; Reyna & Farley, 2006).

The main difference between fuzzy-trace theory and traditional approaches to risk reduction is that the former stresses simple, bottom-line meaning (i.e., gist-based intuition) rather than precise analysis of details. Specifically, the core assumption of expectancy-value approaches is that risks and rewards ought to be analyzed, and that magnitude of rewards should be traded off against magnitude of risks. However, despite the fact that adolescents often overestimate risks, for most adolescents in domains such as sexual risk taking, perceived rewards outweigh perceived risks. Trading off risks and rewards, then, results in a "rational" calculus of risk promotion (i.e., rational from a traditional economics perspective, which differs from rationality in fuzzy-trace theory; Reyna & Farley, 2006). Thus, the low probability of infection with the human immunodeficiency virus (HIV) along with high perceived benefits from sexual activity, rationalize unprotected sex (Adam & Reyna, 2005; Gray et al., 2001; Reyna & Adam, 2003).

According to fuzzy-trace theory, adults do not take fewer risks than adolescents because they think more precisely or elaborately, but, instead, because they think more simply (e.g., Reyna & Brainerd, 2011; Reyna, Estrada et al., 2011). Rather than trade off potentially catastrophic risks against rewards, healthy adults reason more categorically: They categorize unprotected sex as risky, for example, as opposed to calculating the odds of infection from different kinds of partners ("known partners are safe partners," Misovich, Fisher, & Fisher, 1996; 1997). Also, healthy adults retrieve relevant knowledge and values quickly and automatically (i.e., without the necessity for conscious reflection), by virtue of experience (e.g., Fujita & Han, 2009). Cues in the context, and the experience level of a decision maker, determine whether specific knowledge and values are retrieved. Fuzzy-trace theory is not a purely cognitive theory because gist representations are shaped by emotion, motivation, and individual differences (e.g., in world view) and because social constructs such as values are central to the theory (Rivers et al., 2008).

As we discuss in more detail below, modern multi-component educational programs have achieved some success in delaying sexual initiation and reducing risky behaviors among adolescents (e.g., Downs, Murray et al., 2004; Griffin, Botvin, & Nichols, 2006; Kirby & Laris, 2009). Therefore, it makes sense to build on this success by adapting an existing curriculum (*Reducing the Risk*, or RTR) using tenets of fuzzy-trace theory, which is the approach taken here. In another arm of the randomized experiment, we evaluate the standard RTR curriculum, providing the first experimental (i.e., randomized assignment) evaluation of that curriculum. More generally, the approach of emphasizing bottom-line gist (and of automatically retrieving values) can be applied to any health message or educational curriculum (Reyna, 2012a; 2012b). Furthermore, by emphasizing gist representations, which are well preserved over long time periods, a gist-enhanced intervention should produce more sustained effects on behavior.

Specific Motivations for the Study: Research on Effectiveness

As we noted, sexual education programs have evolved from an emphasis on information, to values clarification and cognitive or communication factors, and, currently, to multi-component programs that incorporate prior components plus skills training (for reviews see Kirby, 2007; Kirby et al. 1991; Reyna et al., 2005). Skills are inculcated through role

playing and experiential learning (e.g., practicing refusal skills to resist social pressure to have sex).

Kirby and Laris (2009) reviewed the impact of these curriculum-based programs for adolescents. Of 55 studies published since 1990 that met criteria for inclusion in the review (e.g., sufficient sample size, strong experimental or quasi-experimental design, at least one follow-up assessment), 16 reported delayed initiation of sex, 11 reported a reduced number of partners, and 15 reported reductions in some other measure of sexual risk taking (see also Kirby et al., 2011). Although the proportions of positive outcomes were low (i.e., 16, 11, or 15 positive outcomes out of 55 studies), not all studies collected data on all outcomes. About two out of every three studies reported an effect on at least one behavior for the entire sample or for a subgroup (e.g., males). Similarly, for psychosocial mediators of behavior (e.g., knowledge, attitudes, self-efficacy), about half of the 55 studies reported positive effects of some kind. Nevertheless, only one third of programs improved two or more behaviors; the remainder did not. Overall, programs produced a moderate number of positive effects, but there is clearly room for substantial improvement (DiClemente et al., 2008).

A common limitation of such programs is that behavior change is demonstrated in the short term, but participants are either not followed over the long term or effects dissipate over time (e.g., Jemmott et al., 2010; Kirby & Laris, 2009). Second, many programs have been evaluated using non-experimental designs (that is, without random assignment). Third, among those studies that collect more than two assessments, separate analyses are often conducted by time point, increasing the number of inferential decisions and therefore inflating type I error rates (such an approach also precludes the ability to formally test for predicted time trends). Analysis of change scores, a commonly used alternative, is not without problems, as change scores between pairs of assessments have more measurement error than raw scores (Cronbach & Furby, 1970). Fourth, as time progresses, longitudinal studies suffer from attrition. Commonly used analytical procedures require highly restrictive assumptions about attrition that are unlikely to hold; making them can result in a wide range of artifacts (most typically biased parameter estimates; for a review, see Jeličić & Lerner, 2009).

The present study addresses each of these limitations by: (a) assessing outcomes at five time points; (b) using an experimental design; (c) analyzing time points simultaneously, controlling for the raw score at baseline (Cronbach & Furby, 1970); and (d) using modern methods of analyzing longitudinal data under less restrictive assumptions about the processes underlying attrition. In addition to building on prior work methodologically, the present study also builds on the multiple components of effective curricula.

Multi-component programs synthesize theoretical approaches, such as the theory of reasoned action or planned behavior, social learning, social inoculation, and cognitive behavioral theory (e.g., the behavioral decision-making framework, Fischhoff, 2008). Components include personalization of risk (e.g., belief that one is subject to risk), knowledge (e.g., what can be done to reduce risk), self-efficacy (e.g., belief that one is able to perform relevant actions to reduce risk), anticipated benefits from performing relevant

actions, and practice applying relevant skills. Broadly analogous to expected utility theory, these approaches emphasize rational decision making about the risks and benefits of sex (for overviews of theories of adolescent risky decision making, see Gibbons, Kingsbury, & Gerrard, 2012; Reyna & Rivers, 2008).

Among these programs, RTR demonstrated a sustained effect on sexual initiation in a sample of 758 adolescents in California, among other positive effects (Kirby et al., 1991). Hubbard, Giese, and Rainey (1998) replicated many of these findings of RTR in a rural sample. In addition to behavior, psychosocial mediators (e.g., attitudes, self-efficacy) are a target of the RTR curriculum, and are assessed separately in this study because each has been associated with significant behavior change in prior work. Successfully used in two populations, RTR is one of three curricula in the present study. Advances in research on memory, judgment, and decision-making (for overviews, see Reyna, 2008; Reyna & Farley, 2006) are applied to RTR in an effort to improve on its effects, creating a second curriculum (RTR+). A third curriculum—the control group—involved the same amount of program time, but did not target sexual behavior.

New Curriculum Based on Fuzzy-Trace Theory

The guiding principles of RTR+ are found in fuzzy-trace theory, which has been applied to adolescent risky decision making in laboratory tasks and in real-life decision making (Mills et al., 2008; Reyna, 1996, 2004; Reyna & Adam, 2003; Reyna & Ellis, 1994; Reyna, Estrada et al., 2011; Reyna & Farley, 2006). According to the theory, risk taking arises from developmental differences in knowledge, representation (verbatim and gist), retrieval (e.g., of values), and inhibition. Inhibition (or cognitive control) corresponds to the ability to inhibit inappropriate thoughts or actions in favor of goal-directed ones, and its use increases from childhood to adulthood (e.g., Crockett, Raffaelli, & Shen, 2006; Galvan, 2012; Reyna, 1995; Reyna & Mills, 2007). During the same period, there are changes in the relative reliance on “verbatim” versus “gist” modes of reasoning.

Throughout development, multiple representations of information are encoded at varying levels of precision. Verbatim representations preserve the “facts” of experience, such as exact probabilities (e.g., pregnancy statistics for different forms of contraception; Reyna & Adam, 2003). Gist representations preserve basic meaning and patterns, such as whether an important risk is present versus absent (categorical distinctions) or low versus high (ordinal distinctions) (e.g., Hans & Reyna, 2011). Moreover, gist representations capture the essential bottom line about what matters to an individual (Reyna, 2008; 2012a). Beginning with an initial focus on trading off exact risks and benefits, most decision makers eventually shift to an all-or-none focus for potentially catastrophic decisions (i.e., placing little weight on benefits when the risks, however small, involve catastrophic outcomes). They also increasingly rely on automatic retrieval of relevant values, based on a few cues in the context of decisions.

Evidence supporting this position has come from a wide range of studies with children, adolescents, and adults (e.g., Brainerd, Reyna, & Ceci, 2008; Kühberger & Tanner, 2010; Mills et al., 2008; Reyna, 1996, 2004; Reyna & Brainerd, 1991, 1994, 1995; Reyna & Ellis,

1994; Reyna, Estrada et al. 2011; Reyna & Farley, 2006; Reyna & Lloyd, 2006; Reyna, Lloyd, & Brainerd, 2003). Specific meaning-based biases in decision making emerge with development (and experience), which is a key prediction of the theory (Reyna & Brainerd, 2011; Reyna, Chick, Corbin, & Hsia, 2013; Reyna & Farley, 2006). Such findings reflect a shift towards a form of rationality characterized by an increasing tendency to rely on simplified representations of information, and away from rationality narrowly characterized as involving compensatory trade-offs between risks and benefits, and, a shift towards qualitative and intuitive processing, and away from quantitative, "analytic" tendencies to rely on the verbatim facts of experience (Reyna, 2013). It is the former sort of mature thinking, which characterizes adults, experts, and groups at lower risk (Lloyd & Reyna, 2009; Mills et al., 2008; Reyna et al., 2013; Reyna & Lloyd, 2006), that was encouraged in the modified RTR program of this study.

Theoretically Motivated Outcomes

The outcomes for this study were based on scores of studies of behavior change (for reviews, see Fishbein, 2008; National Institutes of Health, 2009; Reyna & Farley, 2006). Measures had both theoretical justification and empirical support. The primary outcomes were behavior and behavioral intentions, which theories characterize as the outputs of distinct psychosocial mediators that were also measured. These psychosocial mediators can be roughly grouped as follows: attitudes toward sex, beliefs about social norms, motivations (motives to have sex and to not have sex, which are not simply complements of one another psychologically), self-efficacy and behavioral control (the latter added in more recent formulations of theory of planned behavior), and gist and verbatim constructs.

Each scale is justified by prior evidence that it taps a separate construct (e.g., it adds unique variance) which explains risky behaviors in adolescents. For example, attitudes are not manifest in behavior if adolescents lack self-efficacy to act on their attitudes; similarly, some adolescents are motivated to become pregnant but are not particularly motivated to have sex (both motives were assessed); and so on. Reviews and meta-analyses were combed to identify and include separate, reliable, and major predictors of sexual behavior in adolescence. Analyzing each scale separately makes it possible to identify theoretically distinct outcomes and to determine whether each intervention changed that outcome. Note, too, that each outcome was tested controlling for a host of variables that might have produced intercorrelations among outcomes, such as age, gender, race/ethnicity, and baseline score on that outcome.

RTR incorporated all of the content of RTR+. RTR was designed as a multicomponent intervention to affect psychosocial mediators (and, ultimately, behavior), but the range of theoretically predicted outcomes that we investigate has never been tested. However, information that is not understood and retained cannot affect long-term behavior. Thus, RTR + was designed to amplify the uptake of curricular information from RTR in the form of durable mental representations of gist. Hence, RTR+ was hypothesized to more efficiently and effectively produce a range of targeted outcomes. Risks per se were not highlighted more in RTR+ as opposed to RTR—and time spent in the two curricula was the same. Instead, the simple bottom line of risks was emphasized more in RTR+ compared to RTR.

In sum, we manipulated gist representations in RTR+ of the same content as taught in RTR; this type of manipulation was applied in every lesson. Thus, the gist features—such as simple meaning and long-term durability—apply to a range of targeted outcomes (e.g., knowledge, beliefs about norms, etc.). However, given this mechanism, it makes sense to expect that global attitudes and beliefs should be affected more than reasons (which are more analytical than intuitive) and self-efficacy skills (which are more procedural than conceptual). Naturally, we also expected that RTR+ (relative to the RTR or control groups) would be more likely to affect gist outcomes.

Method

Participants

Participants aged 14–19 were recruited from high schools and local youth organizations in Southern Arizona, Northern Texas, and Central New York; 734 adolescents participated in the full intervention and follow-up surveys. Table 1 provides distributional information on sociodemographic and other variables of interest for the three groups. Hispanics were overwhelmingly Mexican-American/Chicano. Spanish versions of all consent materials were available for Hispanic participants, and back-translation was performed on the consent and assent forms to ensure consistency between the language versions. Additional details on the measures in Table 1 (e.g., response categories) are provided in the Appendix.

To minimize attrition, participants were paid a graduated amount, with the largest follow-up payment for completion of the last (12-month) survey. For completion of the presurvey, the full intervention (16 hours), and the postsurvey (which occurred immediately after the intervention), participants were paid a total of $75. Participants then received $15 for completing the 3 month follow-up, $30 for the 6 month follow-up, and $45 for the 12 month follow-up.

A flowchart of participant progress throughout the study can be found in Figure 1. Of 837 initially contacted participants, 87.7% completed the intervention and the immediate follow-up assessment. Of 734 participants who completed the intervention and the immediate follow-up assessment, 80.8% completed a follow-up survey. (Note that Figure 1 reports attendance at each follow-up, but some subjects who could not be scheduled for an earlier follow-up session within our time window completed later sessions and so were not lost to follow-up.) Although fairly prevalent in clinical trial research, current recommendations (Consolidated Standards, 2010) advise against the use of statistical methods that only use data from participants who never dropped out of the study (e.g., complete-case analyses; see discussion below); consequently, we use all data from eligible participants. However, to compare with methods of attrition reporting for complete-case analyses, the total number of eligible participants (734) to complete the final assessment at 12 months was 450 (61.3%). This corresponds to a per-assessment attrition rate of 15% if calculated over the three assessments after the postsurvey (734 participants), and to a rate of 14% if calculated over the four assessments following the presurvey (837 participants). Both rates are within the 10–20% range typically considered acceptable for randomized clinical trials. Differential attrition was not observed: The number of participants to drop out of the study before the final assessment did not vary by intervention.

Participants were randomly assigned to the three interventions and did not differ in baseline characteristics prior to the interventions by more than chance levels (Table 1). That is, of the 47 measures listed, only 2 (4.3%) differed significantly across interventions, roughly what would be expected by chance alone. Those measures were age – participants assigned to RTR+ were slightly younger than participants assigned to RTR [$F(2, 732) = 3.5$, $p = .03$] – and knowledge – participants assigned to RTR+ had slightly lower knowledge scores than participants assigned to RTR [$F(2, 732) = 3.7$, $p = .02$].

Materials

Interventions—RTR is a multi-component intervention, grounded in social learning, social inoculation, and cognitive behavior theories, that emphasizes abstinence (delaying initiation of sex), as well as prophylactic measures to reduce risk. (Several components reflect multiple approaches, such as training in decision-making and communication skills). Participants practice behavioral skills through role-playing and homework activities. For example, one role-playing activity in RTR involved addressing misconceptions about the birth control pill. Following a class discussion about such misconceptions, two volunteer participants read a partially scripted, casual conversation between two hypothetical friends. One friend's statements were entirely scripted and included misconceptions about the pill, such as "the pill will make you fat." The other participant's lines were not scripted: Instead, the participant responded to those misconceptions using points covered during the preceding class. Following the role-playing activity, the health educator led a discussion on how the misconceptions were addressed and alternative ways they could have been addressed. Another example of behavioral skills training included a homework assignment that followed a class discussion of purchasing condoms. Participants visited or called a local store that sold condoms, asked questions about key attributes discussed in class (e.g., latex vs. animal skin, spermicidal vs. non-spermicidal), and reported what they learned at the next class, at which the health educator led a discussion.

Social inoculation theory's emphasis on recognizing social pressure and having the motivation and capability to resist that pressure is also incorporated in class discussions and activities in RTR. Lessons and activities targeted participants' ability to recognize "warning signals" that suggest unsafe sex may be imminent (e.g., being alone with a significant other, lights low and soft music playing, the presence of alcohol), and lessons covering multiple ways to "say no" to sex were practiced in role-playing scenarios. For example, following a discussion of refusal tactics–ways to get out of a situation in which risky sex is imminent–participants role-played a partially scripted scenario in which they applied verbal tactics (discussed in the previous lesson) to avoid sexual advances (e.g., suggesting getting something to eat). The lesson and follow-up discussion addressed salient concerns of dating teenagers, such as tactful ways to defuse such situations without offending a partner. Lessons were written to accommodate same-sex partners (e.g., use of gender-neutral names such as "Lee").

Finally, the personalization of information about sexual risk, a key component of cognitive behavior theory, is also incorporated in RTR. For example, one activity involved simulating pregnancy probability: Given the probability of becoming pregnant for one act of

unprotected sex, and assuming one such act per month, participants in the class drew cards from a hat representing whether they became pregnant. Participants who "become pregnant" stand, and the activity continues for one simulated year (12 draws), at which point the entire class is typically standing. In other words, the activity was structured so that by the end of the simulated year, virtually all of the participants became pregnant (or had gotten someone pregnant). The activity was accompanied by an interactive class discussion about when exactly they became pregnant in the exercise, when they would have the baby, and what the pregnancy would mean to the participant. For example, participants discuss how the pregnancy and baby would affect their own plans for the future (such as going to college) and how it would affect their life in the short-term (such as involvement in extracurricular activities). Comparable discussions and activities were presented for STI infections.

RTR+ is an adapted version of RTR and, therefore, shares its content (i.e., all of the facts covered in RTR+ were also covered in RTR). Like RTR, it emphasizes abstinence as an option to *eliminate* risk in addition to prophylactic measures to *reduce* risk. The key difference between the two interventions is that RTR+ emphasizes "framing" typical sexual decisions in categorical ways that should promote risk avoidance, according to research on fuzzy-trace theory (e.g., Reyna et al., 2011). Specifically, RTR+ accomplishes this by promoting gist extraction (i.e., understanding the essence or bottom line of lessons), automatic retrieval of relevant values, and automatic application of values to gist representations. In RTR+ (but not RTR), short "bottom-line" summaries of important points of the day's lesson were provided at the end of each class, designed to facilitate the encoding and long-term retention of core information.

Thus, the interventions differed with respect to how risk information was mainly characterized, as detailed, quantitative, and verbatim versus qualitative, categorical, and gist-based. For example, RTR included a discussion of quantitative probabilities of various consequences associated with single unprotected and protected sexual acts. RTR+ had a similar discussion, but it also expanded on the numbers by emphasizing the cumulative probability that such a consequence would occur—that many consequences were essentially categorical—they would happen (i. e., they were virtually certain if the risky behavior were engaged in repeatedly over time). By emphasizing the risk of the behavior *in the limit*, focus was redirected from the quantitative value of the risk (e.g., 90% of pregnancy after a year of unprotected sex), which carries little meaning for both adolescents and adults, to a categorical contrast between an event occurring versus not occurring (all-or-none).

For both RTR and RTR+, the fact that many risks were low in probability was clearly communicated (e.g., If sex is unprotected, "A couple has a 1/12 chance of pregnancy each time they have sex."). Both groups participated in an exercise designed to simulate numerical outcomes of repeatedly engaging in unprotected sex over a year, but the categorical point that eventually this activity would result in pregnancy was summarized in RTR+: "Even low risks add up to 100% if you keep doing it." Gist representations did not replace verbatim information about exact risks, but gist emphasized the pragmatic bottom line. Pregnancy might occur the first month, at 6 months, or at 13 months, but the gist is that it "would happen" in about a year. To take another example, in RTR+, the low probability of HIV infection was communicated, but also the gist that "It only takes once to get HIV" and

that virus infections (herpes, HIV, etc.) are categorically different (they can never be cured), so it is not that germs are germs.

Two other important stages involve encouraging storage of relevant risk-avoidant values and facilitating participants' ability to retrieve those values at the time a decision is made. For example, RTR+ included an ongoing activity in which participants frequently updated a checklist of relevant values they personally endorsed. A long menu of options for the checklist was provided in an early session, and as the classes progressed, participants were encouraged to revise their checklist as they reconsidered the values that were important to them. Values were expressed simply (e.g., "Avoid risk") and endorsements were requested in a similarly categorical manner (i.e., endorse it or not), in line with keeping memorial representations of those values as simple and "gist-based" as possible. RTR+ health educators were trained to find ways to relate discussions of sexual risks in any session to values that were important to adolescents and to encourage interactive discussions of those values. The ongoing nature of the activity (continually relating presented information and participants' comments back to those values) provided repeated "reminders" of those values, which helped to cement and strengthen their storage in memory, thus facilitating retrieval of the values—and application to relevant risk representations—in the months following the intervention. However, RTR+ did not attempt to manipulate, or change the endorsement of, participants' values (i.e., gist principles).

Since RTR+ involved providing the information in RTR in addition to summarizing that information, to keep the interventions at the same total duration, certain parts of RTR had to be cut in RTR+. All material that was cut in the creation of RTR+ was in keeping with principles that guided its formulation. For example, in an HIV-risk activity in which students categorized behaviors under green, yellow, and red lights (representing no, medium, and high risk), discussion of less relevant behaviors such as breast feeding (which appears in RTR) were cut from RTR+. Likewise, in class discussions about mistakes that can be made in using condoms, instead of discussing all the possible mistakes that could be made, RTR+ emphasizes the five most common and dangerous mistakes. When information needed to be removed, things that were deemed to have a superficial relevance to the topic at hand were removed.

The control group received a curriculum to improve participants' communication skills and contained no discussion of issues pertaining to adolescent sexuality. Examples of topics relevant to effective communication in adolescence included how to communicate displeasure without putting someone on the defensive, bullying and how to deal with it, and characteristics that are valued in a friend. For example, one class discussion focused on "you" versus "I" methods of communicating displeasure. "You" methods are accusatory and focus on directing blame towards someone else, whereas "I" methods emphasize how someone else's behavior affects a person's feelings. A session on bullying involved the presentation of a taxonomy of bullying types, including a discussion of which types were more common in males versus females. Following this lesson, students participated in interactive discussions in which they reflected on the taxonomy and discussed hypothetical bullying scenarios and how they would deal with them. The control curriculum contained a

comparable number of interactive activities as RTR and RTR+, and its total duration was the same as RTR and RTR+ (16 hours).

Behavioral measures—Primary outcome measures fall into two major groups: self-reported sexual behavior (defined as oral, anal, or vaginal sex) and prophylactic behavior (see below). During data cleaning, 30 of 734 participants were found to have incoherent response patterns to questions about sexual behavior and were omitted from the sample. For sexual behavior, the first variable was initiation status for participants who were sexually abstinent at presurvey (i.e., whether a participant first initiated sexual activity during the study). The second sexual behavior was cumulative number of sexual partners reported over the course of the study.

For prophylactic behavior, three measures of self-reported prophylaxis were analyzed. First, a prophylactic risk index (PRI) was computed by taking the ratio of two quantities covering behaviors during the previous three months: the frequency of prophylactic behaviors (defined as use of a condom or a condom plus foam) over the frequency of sexual encounters (oral, anal, or vaginal). Because they are at least as risk averse as participants who used prophylaxis on all occasions of sexual activity, participants who did not engage in sexual activity in the last three months were assigned the highest score of one on the PRI. The second measure of prophylaxis was whether the participant used a condom the last time they had sex. Participants who remained sexually abstinent throughout the study were excluded from this analysis. Finally, the third measure of prophylaxis was the cumulative number of unprotected sexual encounters, calculated by subtracting the frequency of prophylactic behaviors from the frequency of sexual encounters (both defined above).

Behavioral intentions—A measure of sexual intentions was constructed from responses to five Likert-type items such as “Do you think you will have sex (or have sex again) during the next year?” Analogously, a prophylactic intentions scale was created from six Likert-type items such as “Do you intend to use a condom (rubber) when you have sex?” For each intentions question (sexual or prophylactic; Table 1), one of five responses was selected ranging from *very unlikely* to *very likely*; responses were averaged for overall intentions scores (separately for sexual and prophylactic scales). Higher scores imply higher intentions to engage in the respective behaviors.

Attitudes—A measure of sexual attitudes was constructed from ratings of three Likert-type items such as “I believe it’s okay for people to have sex with a steady boy/girlfriend,” and a prophylactic attitudes scale was created from 24 Likert-type items such as “I believe condoms should always used if a person my age has sex, even if the two people know each other very well.” Items for both measures were rated on a 5-point scale from *strongly disagree* to *strongly agree* and responses were averaged. Higher scores imply more favorable attitudes towards sex and more favorable attitudes towards the use of prophylaxis.

Norms—Three scales composed of aggregated normative Likert-type items were constructed. A measure of perceived sexual norms was constructed from ratings of six items such as “Most of my friends have not had sex yet” (reverse scored) and “Most of my friends believe it’s okay for people my age to have sex with a steady boyfriend or girlfriend.” Items

included both descriptive (beliefs about what peers do) and injunctive (beliefs about what peers should do) content. Items were rated on a 5-point scale from *strongly disagree* to *strongly agree*, and responses were averaged. A measure of perceived parental sexual norms was constructed from four items, such as "How would your mother feel about your having sex at this time in your life?" Items included only injunctive content. Items were rated on a 5-point scale from *strongly disapprove* to *strongly approve*, and responses were averaged. For prophylaxis, a measure of perceived prophylactic norms was constructed from ratings of eight Likert-type items such as "Most adults who are important to me believe condoms (rubbers) should always be used if a person my age has sex, even if the two people know each other very well," and "Most of my friends believe condoms (rubbers) should always be used if a person my age has sex, even if the girl uses birth control pills." Items included only injunctive content. Again, items were rated on a 5-point scale from *strongly disagree* to *strongly agree*, and responses were averaged. For all normative scales, higher scores imply more permissive/favorable norms.

Motives—A scale representing motives for getting pregnant (Unger, Molina, & Teran, 2000) was created from 17 Likert-type items such as "It wouldn't be all that bad at this time in my life [if I had a baby]." Items were rated on a 5-point scale from *strongly disagree* to *strongly agree* and averaged, with higher scores representing a stronger motivation to become pregnant. Participants also rated the extent to which they agreed with 20 statements about the reasons to have–or to not have–sex, such as "I do not want to be a teen parent" and "I feel mature enough to make this decision," on 5-point scales ranging from *strongly disagree* to *strongly agree*. Items were averaged separately for reasons for and against sex, resulting in two 10-item scales. Higher scores on both scales imply more agreement with reasons for or against having sex.

Self-efficacy and perceived behavioral control—A measure of self-efficacy regarding the ability to "say no" to sex was constructed from ratings of four Likert-type items, such as "I feel comfortable refusing to have sex." A measure of self-efficacy in using prophylaxis was constructed from ratings of six Likert-type items, such as "I could succeed in using a condom (rubber) when I have sex." A measure of perceived behavioral control was constructed from ratings of five Likert-type items, such as "It is easy for me to get birth control." All items on these three measures were Likert-type ranging from *strongly disagree* to *strongly agree* that were averaged. Higher scores imply higher self-efficacy/perceived control.

Verbatim and gist measures of risks and benefits perceptions—Gist measures were more general and had less precise responses, relative to verbatim measures, cuing different mental representations, as shown in prior work (see Mills et al., 2008; Reyna, 2008; Reyna et al., 2011). Gist-based perceptions were assessed in five ways. The first scale (categorical risk) contained nine items that measured categorical thinking about risk (e.g., "Even low risks happen to someone"), rated on 5-point scales ranging from *strongly disagree* to *strongly agree* and averaged. Higher scores imply more categorical thinking about risk. The second scale (gist principles) contained 15 simple principles (e.g., "Avoid risk" and "Better to be safe than sorry") that participants endorsed (or not) by checking off

all items that applied to them; the number of endorsements was summed, and higher scores imply a greater tendency to endorse gist-based principles of sexual risk avoidance. The third gist measure (global risk) asked participants to categorize the personal risk of having sex as none, low, medium, or high. Similarly, a global benefit measure asked participants to categorize the personal benefit of having sex on the same rating scale (i.e., from none to high). The final gist measure assessed recognition of danger (i.e., "red light" situations). Participants provided ratings to three Likert-type items assessing the extent to which the participant recognized common warning signals of sex, such as "Being pressured or controlled in any way is a warning signal for unwanted sex," on 5-point scales ranging from *strongly disagree* to *strongly agree*, which were averaged. Higher scores imply greater recognition of warning signals.

Verbatim-based risk perception was assessed with two measures (see Mills et al., 2008; Reyna, 2008; Reyna et al., 2011). First, a specific-risks scale contained five items that mentioned concrete consequences of risky sexual behavior (e.g., pregnancy, HIV-AIDS) and required personal risk estimates of those consequences. Items were rated on 5-point scales ranging from *strongly disagree* to *strongly agree* and averaged, with higher scores implying greater perception of specific risk. A second verbatim-based measure of risk perception required participants to quantify their risk of having a sexually transmitted disease (STD) on a subjective probability scale from 0 to 100 (for which 0 was defined as no chance at all; 50 was as likely as not; and 100 was absolutely certain). Unlike most of the other measures, higher scores on both verbatim measures typically reflect higher risk-taking behavior, and hence higher perceived personal risk (Brewer, Weinstein, Cuite, & Herrington, 2004; Bruine de Bruin, Downs, Murray, & Fischhoff, 2010; Mills et al., 2008).

Knowledge—Participants provided ratings of 27 items assessing knowledge about prophylaxis, sexual risk taking, pregnancy, and sexually transmitted infections (e.g., "Latex condoms prevent HIV better than animal skin condoms") on 5-point scales ranging from *It is false* to *It is true*, which were averaged.

Social desirability—A short form of the Marlowe-Crowne social desirability index (Reynolds, 1982) was also administered to detect biased responders. *True* or *false* responses were provided to 13 items ("No matter who I'm talking to, I'm always a good listener"). Scores were highly similar across groups and, hence, could not account for differences in outcomes.

Procedures

Peer health educator training—The interventions were administered by undergraduate or graduate research assistants (RAs). Using educators only a few years older than participants established rapport, which facilitated class discussions of sensitive topics. Prior to training, RAs were randomly assigned to deliver one of the two treatment interventions (RTR or RTR+) or the control intervention, but to avoid contaminating the delivery of either intervention with material from the other, RAs were never assigned to both RTR and RTR+. Some RA's assigned to one of the two treatment interventions were also cross-trained in the

control curriculum. A total of 25 RAs delivered at least one of the three interventions to participants.

Each RA received over 16 hours of training per intervention (beyond standard human subjects' protection training). First, training began by giving the trainee a notebook containing a complete copy of the curriculum (including detailed lesson plans and learning objectives for each class/module), timelines for discussion topics in each class session, protocols for managing their interaction with participants, access to audio recordings of lessons delivered by trained health educators to actual classes, and (for the two treatment interventions) literature on topics covered in the classes (e.g., covering different methods of prophylaxis and types of sexually transmitted infections). Trainees used this notebook to study the material and reviewed the audio recordings.

After becoming familiar with the curriculum, trainees sat in on lessons with other peer health educators, allowing them to see how more advanced educators were handling common questions, managing time, and directing class discussions. Then, trainees were required to deliver each session that would be administered to participants to mock participants (consisting of other undergraduate and/or graduate students) to a criterion level of performance. This criterion was determined by monitoring a checklist of covered material to ensure that all content was covered, and by judging if the trainee demonstrated an ability to appropriately handle the typical classroom dynamics of a group of high school students (including disruptive behaviors). For each session, trainees received feedback on their performance and repeated the session at other times until criterion performance was achieved. Throughout training and delivery of actual classes to participants, peer health educators consulted with each other and with senior research personnel to ensure new questions and discussion topics that participants raised during classes were handled correctly (e.g., consistent with the content of the curricula). Because RAs became highly familiar with the content and goals of the curriculum to which they were assigned, they cannot be considered "blinded" to the hypotheses.

Randomization—Participants were informed prior to enrollment that they would be randomized to one of three programs. After parental consent and participant assent were received and once a sufficient number of volunteers had been recruited for classes at a site, computer-generated random numbers were used by research personnel (excluding health educators) to allocate enrolled participants to one of the three interventions. (To ensure a balance of interactivity and participation/individual attention, a minimum of five participants and a maximum of 15 were enrolled in any set of 16-hour classes.) For any participants who enrolled at the time of the presurvey (they did not enroll previously but arrived on the scheduled date with all consent materials signed), allocation was determined by drawing random numbers from a bag. While participants were completing the presurvey, health educators were notified of intervention assignments and provided with class rosters, consistent with the goal of allocation concealment during enrollment (Consolidated Standards, 2010). No participant was refused enrollment in the study (except for participants who arrived without signed consent forms): Consent materials distributed to potential participants from target populations (e.g., high schools) contained all information about inclusion criteria, and participants effectively screened themselves prior to enrollment.

During class sessions, health educators took attendance to ensure that random assignments were maintained and that dosage effects could be eliminated by scheduling makeup sessions with students who missed a class or a portion of a class. School personnel were not involved in the random assignment process. Final numbers across interventions are uneven for two reasons. First, an error was made and an allocation of 46 participants was not properly minimized (assignment was random but no constraints based on the current group sizes were placed on the assignments). Second, two health educators trained in the control intervention were unable to continue working on the project at a time when other control educators were not available. Rather than divert resources to training additional educators and under-enroll participants, the decision was made to distribute those participants randomly across the two treatment conditions. Two considerations suggest that these events are not threats to the validity of the random assignment: 1) Educators were randomly assigned to interventions to begin with, making the loss of the control educators a random occurrence, and 2) as illustrated in Table 1, participants in each group were similar across a wide range of variables at presurvey, indicating that random assignment was effective.

Class sessions—Due to scheduling constraints, there was variability in how the 16 contact hours for each curriculum were scheduled. A typical short schedule would involve meeting for two hours each on eight weekdays, after school hours, within a period of two weeks. A typical long schedule would involve meeting for two hours each on three weekdays, after school hours, for three weeks. The average duration during which the 16-hour interventions were delivered was 15.2 days; 83.7% of participants received their full 16-hour intervention in 21 (or fewer) days. Participants were then allotted as much time as they needed to complete the surveys, which consisted of 314 sociodemographic, psychosocial, and behavioral questions (typically taking less than 90 minutes).

Intervention fidelity was monitored in three ways. First, when administering the intervention, educators used the same checklists they were trained with. The checklists included curriculum content (e.g., discussion topics) and amounts of time to spend discussing each topic. As each item was covered, it was checked off. Educators were therefore able to consult the checklist at the end of a session to ensure that all content was covered, and if anything was omitted, they covered it at that time. Second, sessions delivered by health educators were recorded on digital audio recorders, and these recordings were monitored by senior research personnel at regular intervals to ensure all content was being covered and to provide feedback to educators, when necessary. Third, educators met periodically with senior research personnel to discuss any new questions or topics of discussion that arose during classes (educators delivered an answer to the class at the next session).

Participants were asked to not discuss the curriculum outside of class, except with their family. Given the sizes of the target schools and the distribution of participants (e.g. across grades), potential contamination across groups was unlikely. Contamination was further diminished by the administration of the curricula outside of the normal class periods, reducing immediate opportunities for discussion with peers. Thus, the potential for contamination was small, but if it existed, it would operate to diminish differences among groups, providing a more conservative estimate of intervention efficacy.

Statistical Analyses

The general modeling strategy involved first estimating main effects of curriculum assignment and time in models that adjusted for baseline score on the outcome at presurvey, age at the start of the study, female gender, race/ethnicity (White non-Hispanic [reference group], African American, and Hispanic), and the site where the curriculum was administered (Arizona, New York, or Texas [reference group]). Next, we added curriculum by time interaction terms to this covariate-adjusted model to assess whether curriculum efficacy varied across time. Finally, we identified potential variation in curriculum efficacy across age, gender, and racial/ethnic subgroups by testing whether these background characteristics moderated curriculum effects on outcome variables. Although they built upon models accepted in the third stage, results of these analyses are discussed separately given that they were not the primary focus of the study.

As discussed, 5% of variables measured at baseline would be expected to differ between conditions based on chance alone; the observed percentage was 4.3%. Recommendations advise against the inclusion of variables as covariates in planned analyses simply because of detected differences at baseline (Altman, 1998; Assmann, Pocock, Enos, & Kasten, 2000). However, because analyses adjusted for baseline scores on the outcomes at presurvey, any preexisting differences in these variables were controlled. Effectiveness was assessed over the entire follow-up (12-month) period. Because the largest effects for many variables should occur immediately post-intervention, and because differences between interventions theoretically imply differential long-term effects, particular attention was directed to efficacy at postsurvey and at the 12-month assessment.

Discrete time survival analysis (DTSA), linear mixed models, and logistic regression were used to assess effects of intervention assignment on outcomes. DTSA was used to assess how the incidence of sexual initiation varied over time for the three treatment groups (Willett & Singer, 1993). The discrete form is typically used when exact point estimates of the time of “death” (or initiation of sexual behavior) are unavailable. For example, to avoid known error associated with retrospective reports of the exact timing of initiation during intervals between testing, participants provided information pertaining to whether they initiated since the last time they were assessed. Such data are interval censored, with four possible initiation periods corresponding to intervals spanning each of the study’s five assessments. In this analysis, time was therefore treated as a four-level categorical predictor with the first interval as the reference category. A total of 435 participants were sexually abstinent at the baseline assessment and constituted the eligible subject pool.

Following contemporary recommendations for the analysis of continuous outcomes over time e.g., CONSORT guidelines; see American Psychological Association, 2010), linear mixed models were used to examine longitudinal intervention effects on continuous outcomes, including aggregated multi-item scales for psychosocial mediators (e.g., knowledge), the prophylactic risk index, number of sexual partners, and the number of unprotected sexual encounters. Relative to other approaches to analysis of longitudinal data (such as repeated measures analysis of covariance), mixed models provide estimates of effects that are valid under more realistic assumptions about how observations are correlated over time and mechanisms underlying the distribution of missing observations (Jeličić &

Lerner, 2009). Because they are implemented in a regression framework, time can be treated as a continuous variable, and by implication, change in outcomes following curriculum delivery can be modeled more precisely.

As noted earlier, curriculum-induced changes in many outcome measures – in particular, those that directly assess curriculum content (e.g., knowledge) – should often be most pronounced immediately following curriculum delivery, when memory for learned material is strongest. As memory for learned material dissipates over subsequent follow-up sessions at three, six, and 12 months, scores on indices of learned material should of course follow suit. These issues complicate attempts to fit linear models to the data, as (a) there is no reason to suspect that the rate at which information is acquired during the two weeks spanning the pre and postsurvey assessments should equal the rate at which information is forgotten (e.g., Brainerd, Reyna, & Howe, 2009), and (b) if time trends are characterized by initial increases followed by decreases, linear models will not accurately reflect this nonlinear trend. Consequently, we treated the presurvey score on the outcome measure as a separate predictor and modeled change over time in the outcome from the postsurvey to the twelve month assessment. In all linear mixed models we present, time was coded in months centered at the postsurvey assessment to facilitate interpretation of model coefficients in interaction models. That is, the time of the postsurvey assessment is coded as zero. When curriculum by time interactions were detected in subsequent tests of fixed effects, we also computed additional linear transformations of coefficients to identify (and formally test) curriculum effects at specific time points (e.g., to determine whether the positive impact of RTR+ on some outcome dissipated by the three-, six-, or 12-month assessment).

Finally, logistic regression was used to model effects of intervention assignment on whether a condom was used at the last sexual encounter reported at the final 12-month assessment. As in previous models, we initially attempted to control for baseline (presurvey) reports on this measure, but this resulted in perfect prediction errors in early model runs. Consequently, we instead controlled for prophylactic intentions at presurvey – a variable that should also capture pre-existing individual differences in prophylactic behavioral tendencies.

Results

Results are presented in two sections. First, our main focus is on the effects of curriculum (both main effects and – when significant – curriculum by time interactions). In this connection, we briefly overview the efficacy of RTR+ relative to the RTR and control groups, followed by specific findings organized by five categories of outcomes: one category of self-reported behaviors and behavioral intentions plus four categories of psychosocial mediators of curriculum effects, namely, attitudes and norms, motives to have sex or get pregnant, self-efficacy and behavioral control, and gist/verbatim constructs. Second, we present results of additional models examining whether age, gender, or race/ethnicity qualified any of the curriculum findings. (Main effects of age, gender, and race/ethnicity are presented in the Appendix.) For ease of exposition when describing many comparisons, we describe effects succinctly in accordance with model parameterizations. For example, an “RTR+ effect” is understood to represent a contrast of RTR+ with the control group and directly corresponds to the parameter estimate for the RTR+ dummy

coded variable. Any non-parameterized comparison obtained through subsequent linear contrasts is always explicitly described (e.g., RTR+ versus RTR). All final models controlled for participants' age at presurvey, gender, race/ethnicity, the state where the intervention was administered, and (unless otherwise noted) the baseline score on the outcome.

Curriculum Efficacy

Overview of outcomes—As an advance organizer, we should make it clear that we compared each intervention group–RTR+ and RTR--to the control group, but also compared RTR+ directly to RTR on each outcome. Among all 26 outcomes that we evaluated (21 psychosocial mediators including sexual and prophylactic intentions, three measures of prophylactic behavior, and two measures of sexual behavior), 19 showed a significant effect of at least one curriculum relative to the control group. For RTR+, significant improvements relative to the control group were seen in 17 outcomes. For RTR, significant improvements relative to the control group were seen in 12 outcomes. Only RTR+ had a significant impact on measures of sexual behavior. Effects of RTR+ were significantly greater than those of RTR for nine variables, but effects of RTR were significantly greater than RTR+ for three variables.

Self-reported behavior and behavioral intentions—We begin with effects on behavior and behavioral intentions because this is the ultimate guiding question of the research, behavior change. Intervention efficacy for sexual behavior was first assessed by analyzing time to initiation using discrete time survival analysis. Raw hazard rates representing the conditional probability of initiating sexual activity in a specific interval (given sexual abstinence in previous intervals) are shown for each experimental group in Table 2. Overall, hazards show an increasing trend for all groups throughout the study, representing an increasing probability of sexual initiation. Hazards were unsurprisingly low during the first interval – which represents the time during which the intervention was administered (on average, two weeks). However, during the final interval, just under one in five participants in the control condition initiated sexual activity, compared to roughly one in ten participants in the modified (RTR+) intervention. The standard intervention (RTR) fell in between these two groups during the final interval.

Odds ratios from five discrete time survival models are shown in Table 3. We fit three main effect models: a baseline model restricted to effects of time, a covariate model that added effects of site, age, ethnicity, and gender, and a "curriculum" model that added the two curriculum dummy codes. Estimates from these models (the first three columns of Table 3) show that overall time trends followed the same general pattern seen in Table 2, with the likelihood of initiating sexual activity increasing at each successive interval. Relative to the first interval (presurvey to postsurvey), this increase was nonsignificant at the second interval (postsurvey to three months), but was significant at later ones. During the third interval (three to six months), the odds of initiating were roughly two times higher than in the first interval, and they were roughly four times higher during the fourth interval (six to 12 months).

Effects of initiating sex emerged over time for RTR+. In the fourth column of Table 3, estimates are presented from a "full" model that added all possible curriculum by time interaction terms. The significant interval 4 by RTR+ interaction term indicates that relative to participants assigned to the control group, the increase in the odds of initiating sexual activity during the fourth, final interval was significantly lower for participants assigned to the RTR+ curriculum. To understand this interaction effect, note first that because curriculum by time interactions are modeled, the "Interval 4" effect in the full model of Table 3 describes the increase in the odds of initiating between the first and final intervals *for the control group*. This odds ratio shows that among control participants, the odds of initiating sex during the fourth interval was 15.31 times greater than during the first interval. Although not directly parameterized in this model, the corresponding odds ratio for RTR+ is easily calculable and was 2.38. This 84% difference in odds ratios was significant and corresponds directly to the obtained interaction effect of .16. This figure means that the reduction in risk for the RTR+ group was 84%. This interaction effect is shown in Figure 2, which contains estimated hazards from the full model. Estimates from a reduced model in the final column of Table 3 show that the interaction remained when nonsignificant interaction terms were dropped.

One variable in this category, condom use at the last sexual encounter, was not affected by assignment to either curriculum and is not discussed further. Parameter estimates from linear mixed model analyses of the remaining (continuous) behavior and behavioral intention outcomes are shown at the top of Table 4. For each outcome, estimates in this table correspond to one of two models. If nested model comparisons revealed no curriculum by time interactions, we present estimates from models restricted to main effects of curriculum and time (adjusting for background covariates). Conversely, if curriculum effects were found to vary by time, we present estimates from (otherwise identical) interaction models that included curriculum by time interactions. Adjusted means for these outcomes appear at the top of Table 5.

Although the number of sexual partners and number of unprotected sexual acts increased over time, the increase in partners was lower in the RTR+ group, compared to controls. To appreciate relative effect sizes, consider the mean number of sexual partners at 12 months post-intervention: Control = 2.49, RTR = 2.21, RTR+ = 2.15 (Table 5). These figures mean that RTR reduces sexual partners by 0.28. Adding fuzzy-trace theory elements (for basically zero marginal cost) increases the effect size by 21% (0.06/0.28).

For sexual and prophylactic intentions, planned comparisons revealed initial effects of curricula on both measures: At postsurvey, intentions to have sex had decreased in RTR+ (relative to control), and intentions to use prophylaxis had increased in both curriculum groups, compared to controls (as shown by the significant RTR+ and RTR terms in these interaction models; Table 4; see also Table 5). However, these differences had dissipated by the three month assessment and remained nonsignificant at subsequent follow-up assessments. In addition, an omnibus test of the curriculum by time interaction terms in these models did not reach significance, suggesting that changes in sexual and prophylactic intentions over time were largely similar across the three experimental groups (for all

remaining curriculum by time interaction models reported in Table 4, this omnibus test of the interaction was significant).

Attitudes, norms, and motives—Results for the remaining categories of outcomes can also be found in Tables 4 and 5. Although attitudes toward sex became more favorable over time, assignment to RTR+ resulted in significantly less favorable attitudes towards sex overall than assignment to the control group (Table 4) or to RTR (linear contrast: χ^2 [1] = 7.06, $p < .01$). These curriculum differences did not vary over time. Assignment to both RTR and RTR+ resulted in more favorable attitudes toward prophylaxis (compared to controls), but tests of interactions revealed that the RTR+ effect weakened over time. Subsequent linear contrasts showed that whereas the RTR effect held through the 12 month assessment, the RTR+ effect held through six months (χ^2 [1] = 20.07, $p < .001$) and was marginally higher at 12 months (χ^2 [1] = 3.53, $p = .06$). RTR did not differ from RTR+ at any time point.

Independently of curriculum assignment, perceived peer and parental norms regarding sex became more permissive over time, but participants assigned to RTR+ perceived less permissive peer norms overall than participants assigned to the control group (Table 4) or to RTR (linear contrast: χ^2 [1] = 4.56, $p < .05$). This difference did not change over time. Perceptions of prophylactic norms (whether peers used contraception) were enhanced in both RTR+ and RTR relative to controls, and again, these curriculum effects did not vary over time.

Although motivations to become pregnant or to have sex did not change over time or differ across curricula, agreement with reasons to not have sex tended to decline over time and were elevated overall by assignment to both RTR and RTR+ curricula relative to the control group. Once again, tests of interactions were not significant, indicating that these curriculum effects did not vary over time.

Self-efficacy and behavioral control—Over time, self-efficacy for sex (the ability to say “no” or to refuse sex) declined, but perceived control of prophylaxis increased (Table 4 and 5). Both RTR+ and RTR groups perceived that they had greater self-efficacy for refusing sex and for prophylaxis (using contraception), relative to the control group. Both curricula groups also reported higher perceived control of prophylaxis compared to controls. None of these significant effects of curricula interacted with time.

Gist/verbatim constructs—Exposure to the RTR+ and RTR curricula resulted in significant, sustained improvements in knowledge scores and in the recognition of warning signals, relative to controls. Although the magnitude of these curriculum effects dissipated over time (as indicated by the significant, negative interaction terms in Table 4), subsequent linear contrasts demonstrated that RTR+ and RTR curriculum effects remained significant through 12 months for both knowledge, (RTR: χ^2 [1] = 16.56, $p < .001$; RTR+: χ^2 [1] = 38.94, $p < .001$) and warning signal recognition (RTR: χ^2 [1] = 15.29, $p < .001$; RTR+: χ^2 [1] = 44.09, $p < .001$). Moreover, for both of these outcomes, assignment to RTR+ produced larger improvements than assignment to RTR. For knowledge, this advantage for RTR+ participants remained through the six month assessment (knowledge: χ^2 [1] = 7.34, $p < .01$)

and remained marginally elevated through 12 months (χ^2 [1] = 3.17, p = .075). Warning signal recognition remained significantly higher in RTR+ than in RTR participants through 12 months (χ^2 [1] = 5.43, p < .05).

Categorical and global risk perception declined slightly over time, but categorical risk perception was significantly increased overall by assignment to both RTR and RTR+ compared to controls. In addition, assignment to RTR+ produced significantly higher categorical risk perception overall than did assignment to RTR (χ^2 [1] = 22.72, p < .001). Endorsement of gist principles, such as "Avoid risk," was significantly elevated in the RTR group (and was nonsignificantly elevated in the RTR+ group; both groups exhibited relatively high variance on this measure). Relative to controls, who showed no significant time trend, these initial elevations in rates of gist principle endorsement declined for each curriculum group over time. Linear contrasts showed that the initial effect of RTR (relative to control) extended through three months (χ^2 [1] = 6.20, p < .05), but was no longer significant at six months (χ^2 [1] = 2.10, p = .146) or beyond. Gist principle endorsement in the RTR group was also significantly higher than in the RTR+ group through three months (χ^2 [1] = 4.54, p < .05), but this effect had dissipated by the six month assessment (χ^2 [1] = 3.53, p = .06). Thus, a significant increase for RTR after the intervention (and the nonsignificant increase for RTR+) converged with the control group over time (no groups differed at 12 months).

Curriculum Efficacy: Moderating Effects of Age, Gender, and Race/Ethnicity

In a final set of analyses, we probed the boundaries of curriculum effects by testing curriculum interactions with age, gender, and ethnicity. (Main effects—differences observed in outcomes across age, gender, and racial/ethnic subgroups—are presented in the Appendix.) Curriculum interactions with age, gender, and ethnicity were added to final models from the previous stage of analysis (i.e., the models presented in Table 4).

Age and gender—Curriculum effects were largely consistent across age and gender, although there were two exceptions. First, age moderated the RTR+ effect on the perceived benefits of sex. Linear contrasts revealed that among younger participants, assignment to RTR+ led to fewer perceived benefits of sex (a desirable effect of the curriculum) than assignment to the control group (14 year olds: χ^2 [1] = 4.20, p < .05) or to RTR (14 year olds: χ^2 [1] = 4.24, p < .05; and 15 year olds: χ^2 [1] = 4.08, p < .05). However, among 18 year olds, benefit perception was higher in the RTR+ group than in the control group (χ^2 [1] = 4.97, p < .05). Second, curriculum effects on quantitative risk perception varied by gender. Although there were no detectable curriculum effects on this outcome among males, females assigned to RTR perceived significantly greater quantitative risk than females assigned to the control group (χ^2 [1] = 6.60, p < .05) or to RTR+ (χ^2 [1] = 7.62, p < .01).

Race/ethnicity—Race/ethnicity moderated curriculum effects on seven of 26 outcomes. Because some of these ethnicity interactions involved models where curriculum effects also varied by time, a large number of contrasts are needed to fully describe the findings. Consequently, we present a verbal summary here.

Among attitudinal and normative constructs, curriculum interactions with race/ethnicity were observed for three outcome measures: prophylactic attitudes, perceived peer prophylactic norms, and perceived parental norms towards sex. Positive effects of RTR+ and RTR on prophylactic attitudes were sustained through 12 months among Hispanics and through six months for White non-Hispanics. Among African Americans however, assignment to RTR had no detectable impact on prophylactic attitudes, and RTR+ assignment led to more favorable prophylactic attitudes at the postsurvey assessment that could not be detected later. For perceived peer prophylactic norms, RTR assignment induced more favorable perceived norms among Whites and Hispanics. Finally, although there were no main effects of curriculum on perceived parental norms regarding sex, tests of interactions with race/ethnicity revealed that Hispanics assigned to RTR+ perceived less permissive parental norms regarding sex than Hispanics assigned to the control group.

Among measures of self-efficacy and behavioral control, race/ethnicity moderated curriculum effects on prophylactic self-efficacy and on perceived behavioral control of prophylaxis. Increases in prophylactic self-efficacy induced by curriculum were significant for White non-Hispanics (although the effect for Hispanics remained marginal), and increases among RTR participants were largest for Hispanics. An advantage of RTR+ over RTR could be verified among White non-Hispanics. For perceived control over prophylaxis, positive overall effects of both RTR+ and RTR were found to apply to both Whites and Hispanics.

Among gist/verbatim constructs, race/ethnicity moderated curriculum effects on knowledge and categorical risk perception. The positive effects of RTR+ and RTR on knowledge were sustained through 12 months for both White non-Hispanics and Hispanics. However, among African Americans the RTR+ effect could be detected through six months and the RTR effect through three months (although it was marginal through six months). Elevated knowledge among RTR+ compared to RTR participants was demonstrated by White non-Hispanics at all time points. For categorical risk perception, overall increases (relative to control) were seen for all of the race/ethnic groups assigned to RTR+. Increases for the RTR group were demonstrated for both Whites and Hispanics, and increases in RTR+ relative to RTR were detected among Whites and African Americans.

Discussion

Constructs derived from fuzzy-trace theory have been used to explain and predict risk taking in laboratory tasks and in real-world decisions (e.g., Fraenkel et al., 2012; Kuhberger & Tanner, 2010; Mills et al., 2008; Reyna, Estrada et al., 2011; Reyna & Farley, 2006). Here, those constructs were used to amplify the effects of a successful sexual risk-reduction curriculum. Despite communicating the same facts about risk, RTR+ had greater impact than RTR by emphasizing the bottom-line gist of each lesson. Overall, effects were enhanced for gist outcomes, such as categorical risk perception, and for scales that tapped global attitudes and beliefs (Reyna, Estrada et al., 2011).

Most explanations of adolescent risk taking posit dual processes, for example, greater impulsivity and sensitivity to rewards during this developmental period, relative to adults

(e.g., Casey, 2013; Reyna, Chapman, Dougherty, & Confrey, 2012; Steinberg, 2008; see Gibbons, Kingsbury, Gerrard, & Wills, 2011). These dual processes are important, but assuming that processes underlying risk taking ultimately mature does not provide immediate options for designing interventions to change adolescent behavior (see Sunstein, 2008).

Fuzzy-trace theory differs from other dual-process models in distinguishing *impulsivity from intuition* and in emphasizing that *advanced cognition is intuitive*. The main difference between intuition and analysis, in this view, is in the mental representations used to process information. Verbatim representations—precise and detailed—support analysis. In contrast, gist-based intuition operates on simple, bottom-line representations of the *meaning* of information or experience. The modified curriculum emphasized such simple, bottom-line representations of each lesson, facilitating advanced cognition that promotes health (e.g., Fraenkel et al., 2012; Reyna et al., 2011; Reyna & Adam, 2003; Reyna & Lloyd, 2006; see also Fukukura, Ferguson, & Fujita, 2012).

The results of this randomized controlled experiment indicate that the modified RTR+ curriculum was effective at reducing sexual risk taking in adolescence, such as delaying initiation of sex and reducing the number of sexual partners. Note that, with respect to the emergence of differences in sexual initiation over time, the average age of the sample was about sixteen years old at the outset of the study and the majority was not sexually active. In addition, sexual initiation increased over time in this sample, as it does in national samples. Therefore, it makes sense that differences in sexual initiation would not be detected until adolescents reached an age at which they generally initiated sexual activity.

The modified curriculum improved on an established, theoretically motivated multi-component intervention endorsed as effective by the Centers for Disease Control and Prevention (i.e., *Reducing the Risk*; Hubbard et al., 1998; Kirbyet al., 1991). That is, assignment to RTR+ produced superior results relative to the control group, for 17 outcomes, whereas assignment to RTR produced superior results for 12 outcomes (both impressive achievements). RTR+ was significantly more effective than RTR for nine of these outcomes. Effects were sustained over the follow-up period for most outcomes, and remained detectable for 12 outcomes more than a year after the intervention, despite decreased sample sizes as delay increased.

Theoretically sensible intervention effects were identified for 19 outcomes, overall, including behaviors and behavioral intentions as well as psychosocial mediators of behavior. Although public health efforts often focus on behaviors, research has shown that affecting psychosocial mediators—such as attitudes, beliefs about social norms, beliefs about reasons to engage (or not) in the behavior, and perceptions of self-efficacy—can produce behavior change (e.g., Fishbein, 2008; Reyna & Rivers, 2008). Thus, influencing these precursors of behavior over long time periods, as accomplished by both interventions, is worthwhile in promoting public health.

Knowledge is a target of both traditional interventions as well as fuzzy-trace theory. Knowledge is necessary, although not sufficient, for behavior change (e.g., Reyna, Nelson,

Han, & Dieckmann, 2009). We grouped effects of knowledge with gist indicators because our knowledge questions required "transfer of training." In other words, knowledge questions generally measured comprehension and retention of information over time, hallmarks of gist processing. Verbatim representations generally become inaccessible after about a week, and, thus, cannot support long-term changes in knowledge (Reyna & Kiernan, 1994). If adolescents had merely memorized facts by rote (verbatim memory), rather than encoding their gist, knowledge scores would not have remained significantly higher for treatment groups over an interval of a year (Brainerd et al., 2009). Therefore, long-term changes in knowledge suggest that RTR+ influenced gist representations of curricular content. Gist representations may reduce risk taking not only because they are better retained, but because they are also more resistant to effects of stress and strong emotion (Rivers et al., 2008), which have been shown to promote risk taking (Jamieson, Koslov, Nock, & Mendes, 2013).

More specifically, the interventions affected constructs identified in fuzzy-trace theory that go beyond traditional theoretical approaches. For example, after exposure to the curriculum, adolescents were more likely to agree with statements such as "It only takes once to get HIV/AIDS" (i.e., categorical risk perception). Agreement with categorical risk statements is as odds with traditional approaches that stress appreciating degrees of risk (e.g., Downs, Bruine de Bruin, Murray, & Fischhoff, 2004; Fischhoff, 2008), but is consistent with simple gist-based thinking about risk (e.g., Reyna et al., 2011). Adolescents in the RTR+ group were also more able to categorize risky situations as low versus high in danger (i.e., recognition of warning signals), an intuition that ordinarily develops with experience (Reyna, 2008; Reyna et al., 2005). RTR+ and RTR exposed participants to the same quantitative information about risks (and RTR+ did not present more extreme information concerning risks), but RTR+ also communicated the qualitative *essence* of the information. Therefore, these results illustrate that gist-based thinking about risk—the kind of thinking associated with lower risk taking—can be taught to adolescents.

In contrast to effects on gist-based measures of risk perception, interventions did not influence "verbatim" measures of risk perception. Precise perceptions of risk as measured on a 0–100% quantitative scale and as agreement with highly specific risk statements (e.g., about getting an STD in the next 6 months) did not vary by group. Prior research has shown that these gist and verbatim measures of risk perception load on orthogonal factors in principal components analyses and are correlated in different ways with risky behavior (see Mills et al., 2008; Reyna, 2008; Reyna et al., 2011). This result is not unexpected because gist-based, rather than verbatim-based, perceptions and thinking were the targets of the RTR + intervention. The target of RTR+ was fuzzy-trace theory's approach to mental representation, specifically, emphasis on categorical (all-or-none) gist representations of risk (e.g., Reyna et al., 2013; Reyna et al., 2011).

Agreement with gist principles (or simple social values) differed after the intervention for treatment groups, relative to controls, but, surprisingly, differences were greater for the RTR than RTR+ group and effects declined over time. Gist principles refer to gist representations of social and moral values, and we did not attempt to change participants' values. Instead, we asked participants to select their own values, and then helped them practice extracting

the gist of the curricular content and mapping that gist onto their pre-existing values. As the results showed, the rate of agreement with the values/principles was high in all three groups before the interventions began. Adolescents in all groups endorsed, on average, more than ten out of the fifteen gist principles as personally relevant to their sexual behavior. Thus, although gist principles were relevant (and predicted behavior), they did not differ across groups.

The interventions also did not produce significant effects on measures of self-reported prophylactic behavior (number of unprotected sexual acts and the PRI), although they influenced attitudes toward prophylaxis, perceived norms for prophylaxis, and behavioral intentions to use prophylaxis (each of which ordinarily predicts prophylactic behavior). Under commonly used methods of operationalizing prophylaxis in studies of sexual risk, this behavior is defined as conditional on sexual activity. In other words, participants who are not sexually active are not assigned a score on the prophylaxis variable. For a primarily sexually abstinent sample (as ours was), this can result in a substantial loss of eligible participants for the analysis and, consequently, a loss of power. Alternatives to subsample analysis are to measure the absolute number of non-prophylactic behaviors (such as the total number of unprotected sexual acts) or to assign a score to sexually abstinent participants that groups the two types of maximally risk avoidant participants--those who are sexually abstinent and those who always use prophylaxis--together (as with PRI). Multiple approaches were used in the present study. The failure to find effects on prophylaxis--despite intervention effects on psychosocial mediators of prophylactic behaviors--suggests that future implementations of the RTR+ curriculum targeting higher risk adolescents may shed more light on the question of whether there are differential improvements for prophylactic behavior.

Although the effects of curriculum were generally robust across gender, racial, and ethnic groups, these effects were not detectable for all subgroups at all time points. In addition, subgroups differed with respect to some psychosocial mediators. For example, African-Americans perceived themselves at higher risk on verbatim-based questions of risk perception (verbatim risk questions tend to correlate positively with behavioral risk taking) and they perceived greater benefits of having sex, compared to White non-Hispanics. The reverse was true for risk-reducing attitudes, norms, self-efficacy, and gist measures, which were endorsed more often by White non-Hispanics than African-Americans. Subsequent research should move beyond describing such subgroup differences to addressing their social and cultural origins.

In sum, results of this randomized experiment support previous recommendations (e.g., Rivers et al., 2008) to incorporate gist-based manipulations into current public health approaches. Relative improvements in outcomes afforded by RTR+ over RTR were substantial (though RTR produced superior results for two out of 26 measures). The long-term effects reported in this study were adjusted for demographic covariates and evaluated over time, increasing confidence in the results. Because our analyses adjusted for scores on the outcome measures at the presurvey, curriculum effects were independent of any pre-existing differences in the outcome. RTR and RTR+ had highly similar curriculum content; all of the facts covered in RTR+ were also covered in RTR. RTR encompassed a number of

"best practices" and psychological theories. However, minor changes–as straightforward as providing a gist-based review of covered material (i.e., emphasizing categorical representations of risk) at the end of a class–were sufficient to produce differences across a wide range of variables important to adolescent sexual risk taking. Overall, these findings highlight that simple, theory-driven manipulations can be effective in improving on evidence-based programs for reducing risky behaviors in adolescents.

Acknowledgments

Preparation of this article was supported in part by grants to the first author from the National Institute of Mental Health (RO1MH-061211), National Institute of Nursing Research (RO1NR014368-01), and the National Science Foundation (BCS-0840111).

References

Abma J, Chandra A, Mosher W, Peterson L, Piccinino L. Fertility, family planning, and women's health: New data from the 1995 National Survey of Family Growth. Vital and Health Statistics. 1997; 23(19):1–114.

Adam MB, Reyna VF. Coherence and correspondence criteria for rationality: experts' estimation of risks of sexually transmitted infections. Journal of Behavioral Decision Making. 2005; 18:169–186.

Altman, DG. Adjustment for covariate imbalance. In: Armitage, P.; Colton, T., editors. Encyclopedia of Biostatistics. Chichester, UK: John Wiley & Sons; 1998. p. 1000-1005.

American Psychological Association. Publication manual of the American Psychological Association. 6. Washington, DC: American Psychological Association; 2010.

Assmann SF, Pocock SJ, Enos LE, Kasten LE. Subgroup analysis and other (mis)uses of baseline data in clinical trials. Lancet. 2000; 355:1064–1069. [PubMed: 10744093]

Bacik JM, Murphy SA, Anthony JC. Drug use prevention data, missed assessments and survival analysis. Multivariate Behavioral Research. 1998; 33:573–588.

Brainerd CJ, Reyna VF, Ceci SJ. Developmental reversals in false memory: A review of data and theory. Psychological Bulletin. 2008; 134(3):343–382.10.1037/0033-2909.134.3.343 [PubMed: 18444700]

Brainerd CJ, Reyna VF, Howe ML. Trichotomous processes in early memory development, aging, and neurocognitive impairment: A unified theory. Psychological Review. 2009; 116(4):783–832.10.1037/a0016963 [PubMed: 19839684]

Brewer NT, Weinstein ND, Cuite CL, Herrington J. Risk perceptions and their relation to risk behavior. Annals of Behavioral Medicine. 2004; 27:125–130. [PubMed: 15026296]

Bruine de Bruin W, Downs JS, Murray PM, Fischhoff B. Can female adolescents tell whether they will test positive for Chlamydia infection? Medical Decision Making. 2010; 30:189–193. [PubMed: 19706879]

Carlin JB, Galati JC, Royston P. A new framework for managing and analyzing multiply imputed data in Stata. Stata Journal. 2008; 8:49–67.

Casey BJ. The teenage brain: An overview. Current Directions in Psychological Science. 2013; 22(2): 80–81.10.1177/0963721413486971

Centers for Disease Control and Prevention. Youth Risk Behavior Surveillance - 2011. MMWR. Surveillance Summaries. 2012; 61(SS–4)

Consolidated Standards of Reporting Trials. 2010. Retrieved from www.consort-statement.org

Crockett LJ, Raffaelli M, Shen YL. Linking self-regulation and risk proneness to risky sexual behavior: Pathways through peer pressure and early substance use. Journal of Research on Adolescence. 2006; 16:503–525.

Cronbach LJ, Furby L. How should we measure change or should we? Psychological Bulletin. 1970; 74:68–80.

DiClemente RJ, Wingwood GM, Crosby RA, Salazar LF, Rose E, Sales JM, Caliendo AM. Prevalence, correlates, and efficacy of selective avoidance as a sexually transmitted disease prevention strategy among African American adolescent females. Archives of Pediatrics & Adolescent Medicine. 2008; 162(1):60–65.10.1001/archpediatrics.2007.5 [PubMed: 18180414]

Downs JS, Bruine de Bruin W, Murray PJ, Fischhoff B. When “it only takes once” fails: Perceived infertility predicts condom use and STI acquisition. Journal of Pediatric and Adolescent Gynecology. 2004; 17(3):224.

Downs JS, Murray PJ, de Bruin WB, Penrose J, Palmgren C, Fischhoff B. Interactive video behavioral intervention to reduce adolescent females’ STD risk: A randomized controlled trial. Social Science & Medicine. 2004; 59:1561–1572. [PubMed: 15279915]

Fischhoff B. Assessing adolescent decision-making competence. Developmental Review. 2008; 28:12–28.

Fishbein M. A reasoned action approach to health promotion. Medical Decision Making. 2008; 28(6): 834–844.10.1177/0272989X08326092 [PubMed: 19015289]

Fraenkel L, Peters E, Charpentier P, Olsen B, Errante L, Schoen RT, Reyna V. Decision tool to improve the quality of care in rheumatoid arthritis. Arthritis Care & Research. 2012; 64(7):977–985.10.1002/acr.21657 [PubMed: 22392766]

Fujita K, Han HA. Moving beyond deliberative control of impulses: The effect of construal levels on evaluative associations in self-control conflicts. Psychological Science. 2009; 20:799–804. [PubMed: 19493322]

Fukukura J, Ferguson MJ, Fujita K. Psychological distance can improve decision making under information overload via gist memory. Journal of Experimental Psychology: General. 2012 Oct 29. Advance online publication. 10.1037/a0030730

Gerrard M, Gibbons FX, Houlihan AE, Stock ML, Pomery EA. A dual process approach to health risk decision-making: The prototype-willingness model. Developmental Review. 2008; 28:29–61.

Gibbons FX, Kinsgbury J, Gerrard M. Social-psychological theories and adolescent health risk behavior. Social and Personality Psychology Compass. 2012; 6(2):170–183.10.1111/j.1751-9004.2011.00412.x

Gibbons FX, Kingsbury JH, Gerrard M, Wills TA. Two ways of thinking about dual processing: A response to Hofmann, Friese, and Weirs (2008). Health Psychology Review. 2011; 5:158–161.

Gray RH, Wawer MJ, Brookmeyer R, Sewankambo NK, Serwadda D, Wabwire-Mangen F, Lutalo T, Li X, vanCott T, Quinn TC. Probability of HIV-1 transmission per coital act in monogamous, heterosexual, HIV-1-discordant couples in Rakai, Uganda. Lancet. 2001; 357(9263):1149–1153. [PubMed: 11323041]

Griffin KW, Botvin GJ, Nichols TR. Effects of a school-based drug abuse prevention program for adolescents on HIV risk behaviors in young adulthood. Prevention Science. 2006; 7:103–112. [PubMed: 16604429]

Hans VP, Reyna VF. To dollars from sense: Qualitative to quantitative translation in jury damage awards. Journal of Empirical Legal Studies. 2011; 8(S1):120–147.10.1111/j.1740-1461.2011.01233.x

Hoffman, SD. By the numbers: The public cost of teen child bearing. Washington, DC: National Campaign to Prevent Teen and Unplanned Pregnancy; 2006. Retrieved from http://www.thenationalcampaign.org/costs/by-the-numbers.aspx

Hubbard BM, Giese ML, Rainey J. A replication study of Reducing the Risk, a theory-based sexuality curriculum for adolescents. Journal of School Health. 1998; 68(6):243–247. [PubMed: 9719998]

IOM (Institute of Medicine) and NRC (National Research Council). Committee on the Science of Adolescence. Washington, DC: The National Academies Press; 2011. The science of adolescent risk-taking: Workshop report. Retrieved from http://www.nap.edu/openbook.php?record_id=12961

Jamieson JP, Koslov K, Nock MK, Mendes WB. Experiencing discrimination increases risk taking. Psychological Science. 2013; 24(2):131–139. [PubMed: 23257767]

Jeličić EP, Lerner RM. Use of missing data methods in longitudinal studies: The persistence of bad practices in developmental psychology. Developmental Psychology. 2009; 45:1195–1199. [PubMed: 19586189]

Jemmott JB, Jemmott LS, Fong GT. Efficacy of a theory-based abstinence-only intervention of over 24 months: a randomized controlled trial with young adolescents. Archives of Pediatrics & Adolescent Medicine. 2010; 164(2):152–9.10.1001/archpediatrics.2009.267 [PubMed: 20124144]

Johns BT, Jones MN, Mewhort DJK. A synchronization account of false recognition. Cognitive Psychology. 2012; 65:486–518. [PubMed: 22884279]

Kirby, D. Emerging Answers 2007: Research Findings on Programs to Reduce Teen Pregnancy and Sexually Transmitted Diseases. Washington, DC: National Campaign to Prevent Teen and Unplanned Pregnancy; 2007.

Kirby D, Barth RP, Leland N, Fetro JV. Reducing the risk: impact of a new curriculum on sexual risk-taking. Family Planning Perspectives. 1991; 23:253–263. [PubMed: 1786806]

Kirby, D.; Coyle, K.; Alton, F.; Rolleri, L.; Robin, L. Reducing adolescent sexual risk: A theoretical guide for developing and adapting curriculum-based programs. Scotts Valley, CA: ETR Associates; 2011.

Kirby D, Laris BA. Effective curriculum-based sex and STD/HIV education programs for adolescents. Child Development Perspecitves. 2009; 3(1):21–29.10.1111/j.1750-8606.2008.00071.x

Kühberger A, Tanner C. Risky choice framing: Task versions and a comparison of prospect theory and fuzzy-trace theory. Journal of Behavioral Decision Making. 2010; 23:314–329.

Lloyd FJ, Reyna VF. Clinical gist and medical education: Connecting the dots. Journal of the American Medical Association. 2009; 302(12):1332–1333.10.1001/jama.2009.1383 [PubMed: 19773570]

Martin, JA.; Hamilton, BE.; Ventura, SJ.; Osterman, MJK.; Wilson, EC.; Mathews, TJ. National Vital Statistics Reports. Vol. 61. Hyattsville, MD: National Center for Health Statistics; 2012. Births: Final data for 2010.

Mills BA, Reyna VF, Estrada SM. Explaining contradictory relations between risk perception and risk taking. Psychological Science. 2008; 19:429–33. [PubMed: 18466401]

Misovich SJ, Fisher JD, Fisher WA. The perceived AIDS-preventive utility of knowing one's partner well: A public health dictum and individuals' risky sexual behavior. Canadian Journal of Human Sexuality. 1996; 5(2):83–90.

Misovich SJ, Fisher JD, Fisher WA. Close relationships and elevated HIV risk behavior: Evidence and possible underlying psychological processes. Review of General Psychology. 1997; 1:72–107.

National Campaign to Prevent Teen and Unplanned Pregnancy. National data. Washington, DC: National Campaign to Prevent Teen Pregnancy; 2013. Retrieved from http://www.thenationalcampaign.org/national-data/default.aspx

National Institutes of Health. Science of behavior change: Meeting summary; Bethesda, MD. June 15–16, 2009; 2009. Published September 15 2009

Reyna, VF. Interference effects in memory and reasoning: A fuzzy-trace theory analysis. In: Dempster, FN.; Brainerd, CJ., editors. Interference and inhibition in cognition. San Diego, CA: Academic Press; 1995. p. 29-59.

Reyna VF. Conceptions of memory development with implications for reasoning and decision making. Annals of Child Development. 1996; 12:87–118.

Reyna VF. How people make decisions that involve risk: A dual process 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. 2008; 28:850–865. [PubMed: 19015287]

Reyna VF. A new intuitionism: Meaning, memory, and development in fuzzy-trace theory. Judgment and Decision Making. 2012a; 7(3):332–359.

Reyna VF. Risk perception and communication in vaccination decisions: A fuzzy-trace theory approach. Vaccine. 2012b; 30:3790–3797.10.1016/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.

Reyna VF, Adam MB. Fuzzy-trace theory, risk communication, and product labeling in sexually transmitted diseases. Risk Analysis. 2003; 23:325–342. [PubMed: 12731817]

Reyna, VF.; Adam, MB.; Poirier, K.; LeCroy, CW.; Brainerd, CJ. Risky decision-making in childhood and adolescence: A fuzzy-trace theory approach. In: Jacobs, J.; Klaczynski, P., editors. The development of children's and adolescents' judgment and decision-making. Mahwah, NJ: Erlbaum; 2005. p. 77-106.

Reyna VF, Brainerd CJ. Fuzzy-trace theory and framing effects in choice. Gist extraction, truncation, and conversion. Journal of Behavioral Decision Making. 1991; 4:249–262.

Reyna, VF.; Brainerd, CJ. The origins of probability judgement: A review of data and theories. In: Wright, G.; Ayton, P., editors. Subjective probability. New York: Wiley; 1994. p. 239-272.

Reyna VF, Brainerd CJ. Fuzzy-trace theory: An interim synthesis. Learning & Individual Differences. 1995; 7:1–75.

Reyna VF, Brainerd CJ. Dual processes in decision making and developmental neuroscience: A fuzzy-trace model. Developmental Review. 2011; 31:180–206.10.1016/j.dr.2011.07.004 [PubMed: 22096268]

Reyna, VF.; Chapman, S.; Dougherty, M.; Confrey, J., editors. The adolescent brain: Learning, reasoning, and decision making. Washington DC: American Psychological Association; 2012.

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. 2013 Advance online publication. 10.1177/0956797613497022

Reyna VF, Ellis SC. Fuzzy-trace theory and framing effects in children's risky decision making. Psychological Science. 1994; 5:275–279.

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):1125–1142.10.1037/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; 7:1–44.

Reyna VF, Kiernan B. Development of gist versus verbatim memory in sentence recognition: Effects of lexical familiarity, semantic content, encoding instructions, and retention interval. Developmental Psychology. 1994; 30(2):178–191.10.1037/0012-1649.30.2.178

Reyna VF, Lloyd FJ. Physician decision-making and cardiac risk: Effects of knowledge, risk perception, risk tolerance, and fuzzy processing. Journal of Experimental Psychology: Applied. 2006; 12(3):179–195.10.1037/1076-898X.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. Interference processes in fuzzy-trace theory: Aging, Alzheimer's disease, and development. In: MacLeod, C.; Gorfein, D., editors. Inhibition in cognition. Washington: APA Press; 2007. p. 185-210.

Reyna VF, Nelson W, Han P, Dieckmann NF. How numeracy influences risk comprehension and medical decision making. Psychological Bulletin. 2009; 135(6):943–973.10.1037/a0017327 [PubMed: 19883143]

Reyna V, Rivers SE. Current theories of risk and decision-making. Developmental Review. 2008; 28:1–11. [PubMed: 19255598]

Reynolds WM. Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. Journal of Clinical Psychology. 1982; 38:119–125.

Rivers SE, Reyna VF, Mills BA. Risk taking under the influence: A fuzzy-trace theory of emotion in adolescence. Developmental Review. 2008; 28:107–144. [PubMed: 19255597]

Royston P. Multiple imputation of missing values: Further update of ice, with an emphasis on interval censoring. Stata Journal. 2007; 7:445–464.

Rubin, DB. Multiple imputation for non-response in surveys. New York: John Wiley & Sons; 1987.

Satterwhite CL, Torrone E, Meites E, Dunne EF, Mahajan R, Ocfemia MCB, Weinstock H. Sexually transmitted infections among U.S. women and men: Prevalence and incidence estimates, 2008. Sexually Transmitted Diseases. 2013; 40(3):187–193. [PubMed: 23403598]

Schafer JL, Olsen MK. Multiple imputation for multivariate missing-data problems: A data analyst's perspective. Multivariate Behavioral Research. 1998; 33:545–571.

Steinberg L. A social neuroscience perspective on adolescent risk-taking. Developmental Review. 2008; 28:78–106. [PubMed: 18509515]

Suellentrop, K. Teen contraceptive use. Washington, DC: The National Campaign to Prevent Teen Pregnancy; 2006.

Sunstein CR. Adolescent risk-taking and social meaning: A commentary. Developmental Review. 2008; 28:145–152.

Unger JB, Molina GB, Teran L. Perceived consequences of teenage childbearing among adolescent girls in an urban sample. Journal of Adolescent Health. 2000; 26:205–212. [PubMed: 10706168]

Weinstock H, Berman S, Cates W. Sexually transmitted diseases among American youth: Incidence and prevalence estimates. Perspectives on Sexual and Reproductive Health. 2004; 36:6–10. [PubMed: 14982671]

Willett JB, Singer JD. Investigating onset, cessation, relapse, and recovery: Why you should, and how you can, use discrete-time survival analysis to examine event occurrence. Journal of Consulting and Clinical Psychology. 1993; 61:952–965. [PubMed: 8113496]

Appendix

Participants

The most frequently reported "other" racial/ethnic group was Asian. These responses were grouped with those of White/Caucasian non-Hispanics in Table 1 and in analyses for two reasons. First, relative to other groups, reports of Asian race/ethnicity was rare. Second, Asians' responses on other variables closely paralleled those of White/Caucasian non-Hispanic participants.

Table 1 Variables

Codings for descriptive variables in Table 1 were as follows: "What kind of grades do you usually get in school?" with "A's," B's," "C's," "D's," and "F's" as response options; "Do you get a free school lunch?" with "No," "Yes," and "Don't know" as response options; "Where do you live right now?" with "I live with both parents," "I live with a single parent," "I live with a parent and step-parent," "I live part time with both families (both parents have custody)," "I live with other relatives (not my parents)," "I live in a group home," "I live with a foster family," and "I live on my own or with friends" as response options (collapsed into a trichotomous "lives with" one, both, or no parents variable); "How important would you say religion is to you? (religiosity) with five response options ranging from "Not at all" to "Very; "In general, how many hours per day are you without any adult supervision?" with "Less than 1 hour," "1–2 hours," "3–4 hours," and "More than 4 hours" as response options; "What is the highest level your [parent] completed in school?" (asked separately for mother and father and averaged for a single parental education measure; "Don't know" responses were treated as missing and the answered item for the other parent – if present – was used for the overall score).

Measures

Consistency

Prior to analyses, standard checks on logical relations between variables were assessed; missing values were substituted when they could be logically inferred and there were no inconsistencies. For example, a participant who missed the 3-month survey but provided a "no" response to the question "Have you ever had sex?" at presurvey, postsurvey, and 6 months was inferred as not having sex in the interval between the postsurvey and 3-month follow-up. Similarly, a participant who said "yes" to the ever-had-sex question at 3 months could not have *initiated* sexual activity between the 3- and 6-month assessments.

Initiation of sexual activity

Initiation status was determined by responses to the yes-no question, "Have you ever had sex?" Status was corroborated by answers to additional, specific yes-no questions: "Have you ever had vaginal sex?," "Have you ever had anal sex?," and "Have you ever had oral sex?"

Number of sexual partners

Participants were asked about the total number of male and female sexual partners they ever had. Responses to these two questions were summed to create an overall number-of-partners variable, and a cumulative version was computed by aggregating it with responses at previous time points.

Prophylactic risk index and number of unprotected sexual encounters

These variables were calculated from questions about the types of prophylaxis used in the last 3 months and the total number of sexual encounters in the last 3 months. The total number of prophylactic behaviors involving a condom in the last 3 months was calculated by summing responses to the questions "If you used contraception in the last three months, how many times did you use condoms (rubbers)?" and "If you used contraception in the last three months, how many times did you use condoms (rubbers) and foam?" The total number of vaginal, oral, or anal sexual episodes in the last 3 months was calculated by summing responses to "In the last three months, I had vaginal (regular) sex_ times," "In the last three months, I had oral sex_ times," and "In the last three months, I had anal (rectal) sex_ times." The ratio of these two numbers (the PRI) was computed to obtain the proportion of protected sexual episodes over the previous 3 months. (Recalculating this ratio without oral sex in the denominator does not change results.) Because they are not more risk seeking than participants who used prophylaxis on all occasions of sexual activity, participants who reported no sexual activity over the last 3 months were assigned the highest score of one on the PRI. The difference between these two variables corresponds to the number of unprotected sexual encounters during the past 3 months. A cumulative version of the variable was calculated by summing responses of the current time point with previous time points.

Condom use at last sexual encounter

Participants checked off one or more of six options to the question “If you have had sex, what method(s) of birth control did you and your partner use to prevent pregnancy the last time you had sex?” The available options were “I have never had sex,” “No method was used,” “Birth control pill,” “Condom (rubber),” “Some other method (ex. Diaphragm, IUD),” and “I am not sure.” Affirmative responses to the condom option at the 12-month assessment were used to construct a dichotomous measure of whether condoms were used at the last sexual encounter.

Statistical Analyses

Missing Data in Discrete Time Survival Analysis

Although survival models properly account for right censored cases (that is, abstinent participants who leave the study and do not return), they do not optimally handle *intermittent* missed response (IMR) patterns (for an overview of this issue for discrete time models, see Bacik, Murphy, and Anthony, 1998). For example, if a participant misses the assessment at time t-1 but reports having initiated sex by the assessment at time t, sexual initiation could have occurred during the interval spanning times t-2 and t-1 or during the interval spanning times t-1 and t. Consequently, prior to fitting this set of models, IMRs were multiply imputed following standard recommendations (Schafer & Olsen, 1998). Stata 11 (StataCorp, 2009) was used to impute 10 datasets (Carlin, Galati, & Royston, 2008; Royston, 2007), simultaneously analyze the data, and produce combined parameter estimates standard errors, confidence intervals, and associated significance tests in accordance with Rubin’s rules (Rubin, 1987). Only 15 (3.4%) of the 435 eligible participants showed an IMR pattern and had data imputed for these analyses. Estimates from analyses of these data are valid under the *missing at random* (MAR) assumption, which is less restrictive than the *missing completely at random* (MCAR) assumption required of techniques such as complete case analysis.

Covariance structure of random effects for linear mixed models

In preliminary runs of linear mixed models, nested models were compared with chi-square difference tests to evaluate the necessity of random slopes (subject-specific variation in time trends), random intercepts (subject-specific variation in starting points), and for models requiring both, optimal covariance structures of those random effects (e.g., independent versus unstructured). The vast majority of outcomes required both random intercepts and random slopes and exhibited an independent covariance structure.

Results: Main Effects of Age, Gender, and Race/Ethnicity

In this section, we present outcome differences across age, gender, and racial/ethnic subgroups. All effects discussed in this section correspond to models discussed in the main text and presented in Tables 4 and 5.

Age

Older participants differed significantly from younger participants on six outcomes. Although they had lower PRI scores (i.e., their condom use behavior put them at greater risk; b = −.03, SE = .01, $p < .01$), they also had greater self-efficacy in refusing sex (b = .05, SE = .02, $p < .05$), perception of categorical risk (b = .05, SE = .01, $p < .01$) and recognition of warning signals (b = .09, SE = .02, $p < .01$). Older participants also reported higher agreement with both reasons to have sex (b = .05, SE = .02, $p < .01$) and with reasons to not have sex (b = .04, SE = .02, $p < .05$).

Gender

Females differed significantly from males on 15 outcome variables. Regarding sex, they reported fewer sexual partners (b = −.36, SE = .14, $p < .01$), parental norms that were less favorable to sex (b = −.14, SE = .04, $p < .001$), lower agreement with reasons to have sex (b = −.09, SE = .04, $p < .05$), higher agreement with reasons to not have sex (b = .12, SE = .03, $p < .001$), and greater self-efficacy in refusing sex (b = .17, SE = .04, $p < .001$). Females' attitudes, beliefs, and intentions were also more favorable to prophylaxis than were males: They had greater intentions to use prophylaxis (b = .12, SE = .04, $p < .01$), more favorable attitudes toward prophylaxis (b = .13, SE = .03, $p < .001$), and more favorable beliefs about prophylactic norms among peers (b = .14, SE = .04, $p < .001$). Females also reported greater self-efficacy in using prophylaxis (b = .19, SE = .04, $p < .001$) and greater perceived control over prophylaxis (b = .15, SE = .04, $p < .001$) than males. Finally, females also reported higher knowledge (b = .06, SE = .03, $p < .05$), perceived greater categorical risk from sex (b = .13, SE = .03, $p < .001$), and endorsed more gist principles (b = .57, SE = .16, $p < .01$) than males, but they perceived fewer benefits of sex (b = −.16, SE = .05, $p < .01$) and lower specific risk (b = −.20, SE = .03, $p < .001$).

Race/ethnicity

African Americans were significantly more likely to initiate sex than White non-Hispanics over the course of the study (81%: see Table 3) and were marginally more likely to use a condom at their last sexual encounter (OR = 1.9, SE = .5, $p = .07$). They were also less likely to agree with reasons to not have sex (b= −.08, SE = .04, $p < .05$) and had lower self-efficacy to refuse sex (b= −.14, SE = .05, $p < .01$). Their prophylactic self-efficacy (b= −.13, SE = .05, $p < .01$) and perceived behavioral control over prophylaxis (b= −.11, SE = .05, $p < .05$) were also lower than White non-Hispanics', and they reported less favorable attitudes towards prophylaxis (b= −.14, SE = .03, $p < .001$) and perceived peer prophylactic norms (b= −.15, SE = .04, $p < .001$). African Americans were also less likely to recognize warning signals (b= −.17, SE = .06, $p < .01$), perceived fewer benefits of sex (b= −.22, SE = .06, $p < .001$), and reported lower knowledge scores (b= −.18, SE = .03, $p < .001$) than White non-Hispanics, and although they perceived lower categorical risk (b= −.10, SE = .04, $p < .01$), they also perceived themselves to have higher specific risk (b= .11, SE = .04, $p < .01$) and quantitative risk (b= 2.8, SE = 1.1, $p < .05$), Although Hispanics and White non-Hispanics did not differ overall on any outcome measure, African Americans differed from Hispanics on three variables: Hispanics endorsed more reasons to not have sex (χ^2 [1] = 5.31, $p < .05$),

and they reported more favorable prophylactic attitudes (χ^2 [1] = 6.89, $p < .05$) and greater knowledge of curriculum material (χ^2 [1] = 6.20, $p < .05$).

Health Educator Effects

Twenty-five individuals were randomly assigned to interventions as health educators or RAs (with constraints noted in the text). Despite random assignment, it is reasonable to wonder about educator (i.e., trainer-level) effects. That is, the educators in RTR (and RTR+) could have been more effective at dissuading young people from performing risky sexual behaviors. Because of the number of educators/trainers (25), overall effects were not localized to a few teachers. Also, clustering error at the trainer-level would be too under-powered as an analytical approach, and, moreover, the level of inference of interest is at the individual-level, not the trainer-level. However, we inspected the results separately by educator/trainer and did not observe substantial variation across educators/trainers, probably due to the stringent criterion of performance required in their preparation (prior to teaching classes), the detailed lesson plans and objectives for each class, and the ongoing monitoring of teaching performance in classes (which was used to catch problems early and remediate them).

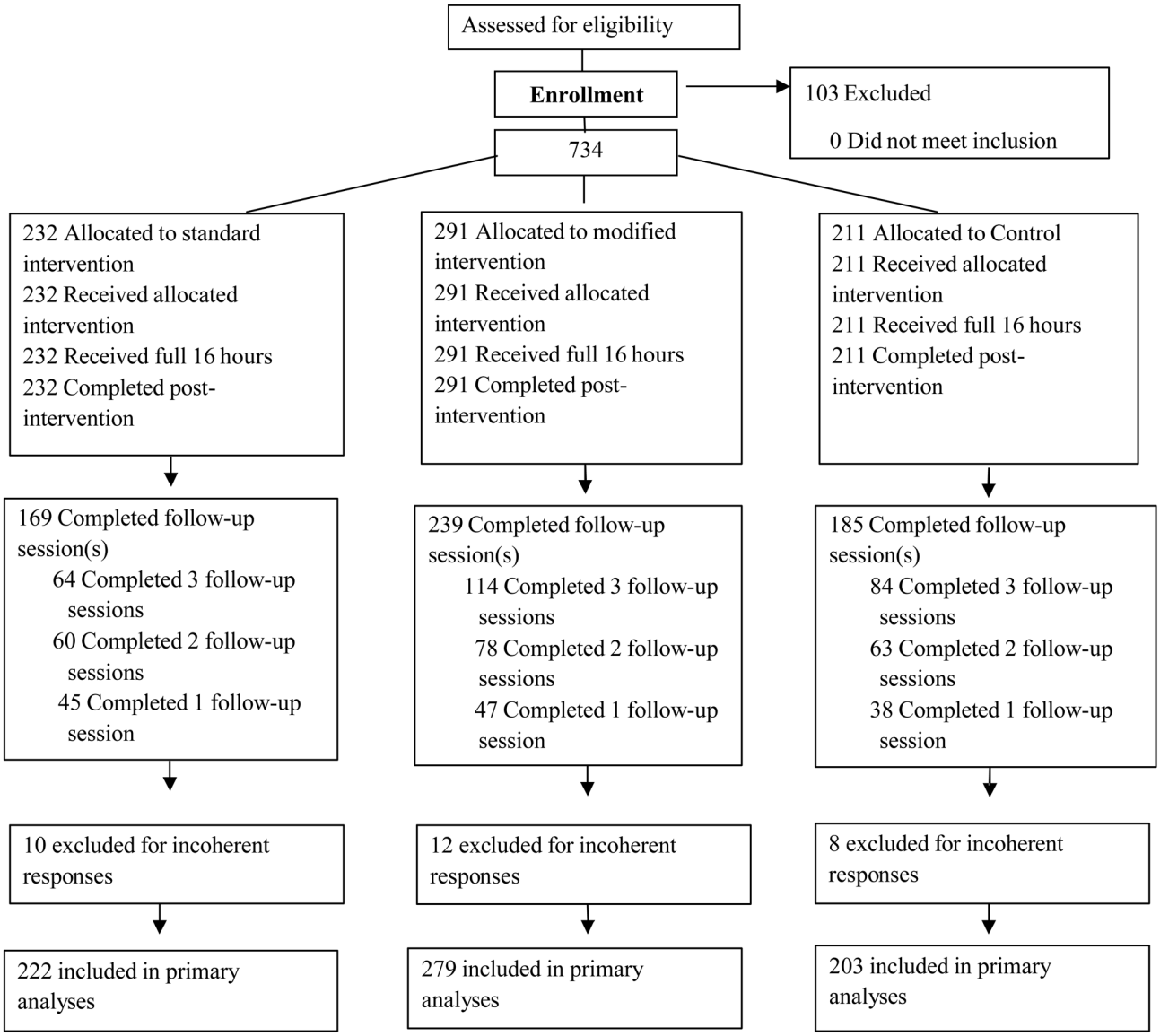

Figure 1.
Flow diagram of the progress through the phases of the parallel randomized experiment of three groups: Enrollment, intervention allocation, follow-up, and data analysis

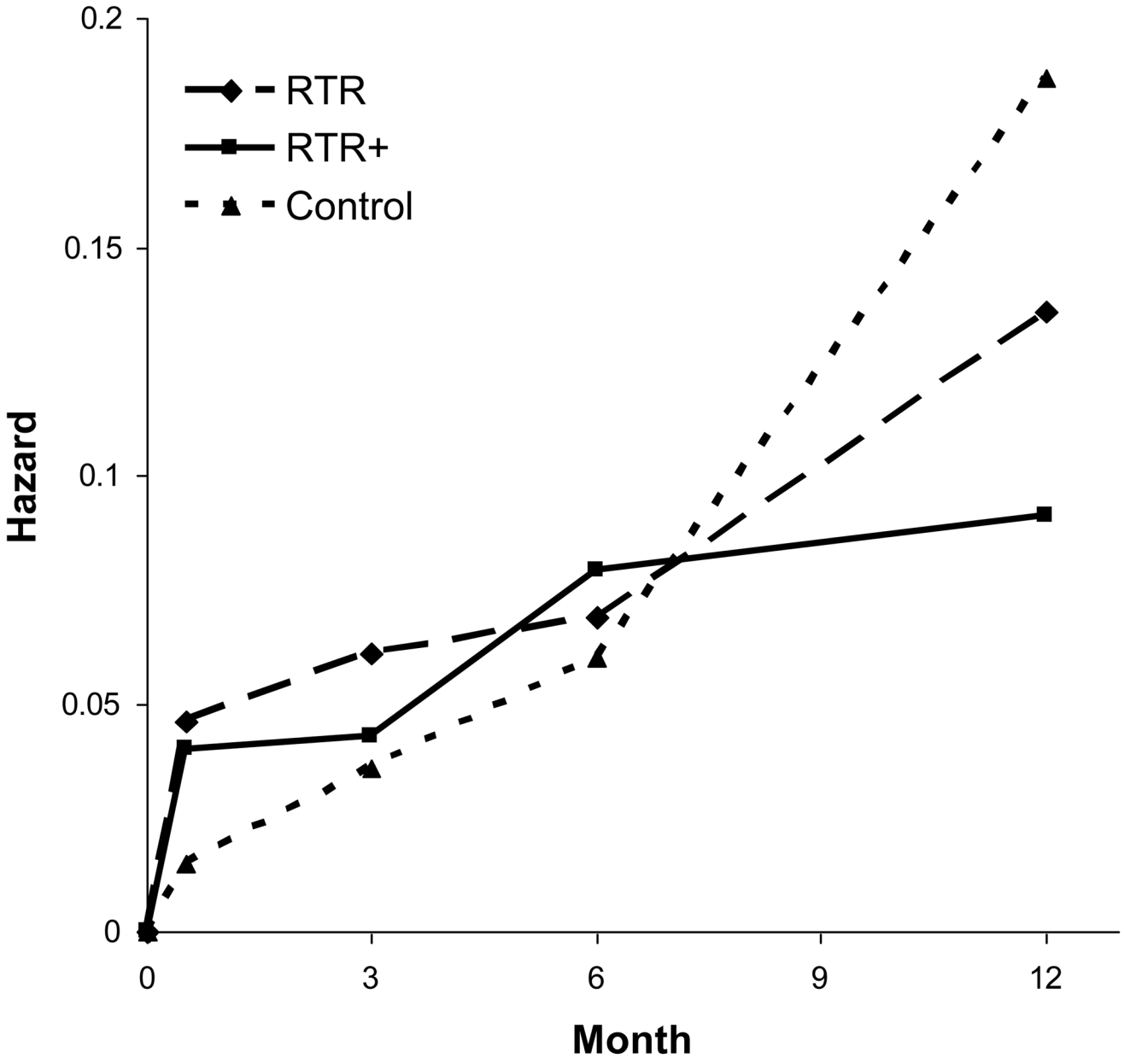

Figure 2.
Sexual initiation: Estimated hazard functions from the full interaction discrete-time survival model by intervention

Table 1

Comparability of Control, RTR, and RTR+ Intervention Groups

Variable	Control (n=211)	RTR (n=232)	RTR+ (n=291)
Sociodemographic variables			
Age, mean (SD)	16.05 (1.1)	16.07 (.99)	15.86 (.99)
Female	117 (55.5)	125 (53.9)	180 (61.9)
Hispanic	33 (15.6)	32 (13.8)	56 (19.2)
African American	57 (27.0)	69 (29.7)	76 (26.1)
Caucasian	105 (49.8)	100 (43.1)	123 (42.3)
Grades in school	1.94 (.86)	1.99 (.83)	2.04 (.87)
Receives free lunch	45 (21.3)	62 (26.7)	87 (29.9)
Hours without adult supervision, mean (SD)	2.99 (1.0)	3.08 (1.0)	2.92 (1.0)
Lives with both parents	111 (52.6)	106 (45.7)	146 (50.2)
Lives with one parent	55 (26.1)	55 (23.7)	72 (24.7)
Parental education, mean,(SD)	2.82 (.83)	2.87 (.91)	2.72 (.93)
Psychosocial mediators, mean (SD)			
Intentions to have sex (α = .92)	2.10 (1.2)	2.34 (1.1)	2.14 (1.2)
Intentions to use Prophylaxis (α = .87)	3.17 (.73)	3.13 (.80)	3.12 (.78)
Sexual attitudes (α = .81)	1.65 (.98)	1.80 (.99)	1.69 (.99)
Prophylactic attitudes (α = .82)	3.01 (.41)	3.01 (.43)	2.98 (.43)
Pregnancy attitudes (α = .79)	1.75 (.55)	1.79 (.57)	1.75 (.58)
Perceived sexual norms (α = .72)	1.79 (.75)	1.88 (.71)	1.82 (.72)
Perceived parental sexual norms (α = .87)	.91 (.87)	.93 (.91)	.78 (.86)
Perceived prophylactic norms (α = .81)	3.10 (.60)	3.10 (.62)	3.08 (.62)
Perceived behavioral control – Prophylaxis (α = .72)	2.85 (.66)	2.80 (.68)	2.81 (.69)
Self-efficacy in using prophylaxis (α = .82)	2.93 (.67)	2.96 (.69)	2.89 (.70)
Self-efficacy in “saying no” to sex (α = .71)	2.91 (.71)	2.79 (.71)	2.83 (.74)
Specific risk perception (α = .82)	.37 (.54)	.36 (.53)	.43 (.59)
Quantitative risk perception	6.33 (14.1)	8.27 (15.9)	10.03 (19.6)
Categorical risk perception (α = .71)	2.94 (.58)	2.90 (.52)	2.92 (.56)
Perceived global benefits of sex	1.32 (.93)	1.38 (.98)	1.26 (.95)
Perceived global risks of sex	1.85 (1.1)	1.83 (1.0)	1.81 (1.1)
Gist principle endorsement (α = .81)	10.79 (3.5)	10.58 (3.3)	11.03 (3.3)
“No risk is better than some risk” endorsement, No. (%)	160 (75.8)	166 (71.6)	220 (75.6)
“Less risk is better than more risk” endorsement, No. (%)	115 (54.5)	120 (51.7)	138 (47.4)
Knowledge (α = .79)	2.82 (.43)	2.87 (.42)	2.77 (.44)
Reasons to have sex (α = .79)	1.71 (.71)	1.74 (.68)	1.70 (.72)
Reasons to not have sex (α = .79)	2.79 (.57)	2.77 (.57)	2.83 (.59)
Recognition of warning signals (α = .50)	1.93 (.76)	1.85 (.72)	1.92 (.78)
Delinquency (α = .79)	.46 (.60)	.48 (.57)	.48 (.54)
Religiosity	3.44 (1.3)	3.47 (1.3)	3.45 (1.3)
Index of peer relations (α = .94)	1.94 (.69)	1.88 (.65)	1.95 (.70)

Variable	Control (n=211)	RTR (n=232)	RTR+ (n=291)
Marlowe-Crowne Soc. Des. index (short form; α = .66)	.46 (.21)	.44 (.21)	.44 (.21)
Sexual/prophylactic behavior			
Had sex in the last 30 days	41 (19.4)	57 (24.6)	50 (17.2)
Ever had sex	88 (41.7)	100 (43.1)	102 (35.1)
Ever had vaginal sex	70 (33.2)	80 (34.5)	86 (29.6)
Ever had anal sex	18 (8.5)	19 (8.2)	19 (6.5)
Ever had oral sex	76 (36.0)	82 (35.3)	84 (28.9)
Number of sexual partners, mean (SD)	1.27 (3.3)	1.29 (3.2)	.88 (2.0)
Currently dating someone	82 (38.9)	92 (39.7)	126 (43.3)
Prophylactic risk index (PRI), mean (SD)	.81 (.36)	.82 (.35)	.85 (.32)
No. unprotected sexual encounters, mean (SD)	2.92 (11.2)	3.06 (9.6)	1.92 (10.9)

Note. Data are expressed as No. (%) of participants unless otherwise noted. α = Cronbach's α for aggregated multi-item measures; Soc.des. = Social desirability.

Ratings of intentions, attitudes, norms, reasons to have sex, reasons not to have sex, self-efficacy, perceived control, categorical risk, warning signals, and specific risk were coded as 0 to 4 with higher scores indicating greater likelihood/agreement/approval; quantitative risk ranged from 0 to 100 with higher scores indicating higher risk; global risk and benefit were coded as 0 (none) to 3 (high); knowledge was coded 0 (false) to 4 (true) and false items were reverse coded; "grades in school" was coded 1 to 5 with lower scores reflecting higher grades; hours without adult supervision and parental education were coded 1 to 4 with higher scores reflecting less supervision and higher education, respectively; delinquency was coded on a 0 to 4 scale with higher scores reflecting more delinquent behavior; religiosity and the index of peer relations were coded from 1 to 5, with higher scores reflecting more religiosity and poorer relations, respectively. Additional details on measures are in the Appendix.

Table 2

Raw hazard rates for initiating sexual activity by intervention

	Intervention		
Interval	**Control**	**RTR**	**RTR+**
1 (Presurvey to Postsurvey)	.016	.054	.044
2 (Postsurvey to 3 months)	.039	.072	.046
3 (3 months to 6 months)	.064	.082	.083
4 (6 months to 12 months)	.189	.159	.095

Note. Values represent the probability of initiating sexual activity given no sexual activity in previous intervals and are not adjusted for demographic background variables.

Table 3

Odds ratios from discrete-time survival analysis of sexual initiation

	Model				
Variable	**Baseline**	**Covariate**	**Curriculum**	**Full**	**Reduced**
Interval 2	1.32	1.36	1.36	2.50	1.37
Interval 3	2.04*	2.13*	2.13*	4.28	2.15
Interval 4	4.11***	4.34***	4.33***	15.31***	9.12***
Arizona		0.81	0.84	0.83	0.83
New York		1.53	1.50	1.52	1.51
Age		1.07	1.06	1.06	1.06
Hispanic		1.29	1.29	1.30	1.30
African-American		1.81*	1.78*	1.81*	1.80*
Female		0.83	0.84	0.84	0.84
RTR			1.16	3.20	1.69
RTR+			0.94	2.75	1.52
Interval 2 X RTR				0.54	
Interval 2 X RTR+				0.43	
Interval 3 X RTR				0.36	
Interval 3 X RTR+				0.49	
Interval 4 X RTR				0.21	0.40
Interval 4 X RTR+				0.16*	0.29*

Note. Reference for interval 2–4: interval 1; reference for Site = Texas; reference for Ethnicity = White/other; reference for curriculum = Control.

* $p<.05$,

*** $p<.001$

Table 4

Unstandardized Parameter Estimates and Standard Errors from Linear Mixed Models Predicting Changes in Behavioral Outcomes and Psychosocial Mediators Over Time

	Predictor				
Outcome	**RTR**	**RTR+**	**Month**	**Month X RTR**	**Month X RTR+**
Behavior and behavioral intentions					
Number of sexual partners	−.284 (.177)	−.345 (.165)*	.073 (.015)***		
Number of unprotected sex acts	−.287 (.584)	−.458 (.542)	.672 (.104)***		
Prophylactic risk index	−.035 (.021)	−.007 (.019)	−.005 (.002)		
Intentions to use prophylaxis	.159 (.060)**	.124 (.056)*	.004 (.005)	.001 (.007)	−.008 (.006)
Attitudes and norms					
Attitudes toward sex	.025 (.049)	−.095 (.045)*	.018 (.003)***		
Attitudes toward prophylaxis	.151 (.035)***	.210 (.032)***	.005 (.003)	−.004 (.004)	−.010 (.004)**
Perceived sexual norms	−.011 (.039)	−.088 (.036)*	.020 (.002)***		
Perceived parental sexual norms	.000 (.051)	−.025 (.047)	.011 (.003)***		
Perceived prophylactic norms	.092 (.047)*	.087 (.043)*	.002 (.002)		
Motives					
Reasons to get pregnant	−.045 (.030)	−.032 (.027)	.001 (.002)		
Reasons to have sex	.025 (.045)	−.031 (.041)	.001 (.003)		
Reasons not to have sex	.143 (.040)***	.167 (.036)***	−.013 (.002)***		
Self-efficacy and behavioral control					
Self efficacy (“saying no” to sex)	.190 (.052)***	.244 (.048)***	−.006 (.003)*		
Prophylactic self-efficacy	.101 (.049)*	.177 (.045)***	.004 (.002)		
Behavioral control (prophylaxis)	.133 (.048)**	.129 (.045)**	.007 (.002)**		
Gist/verbatim constructs					
Knowledge	.412 (.049)***	.549 (.046)***	.008 (.003)**	−.015 (.004)***	−.017 (.004)***
Categorical risk perception	.104 (.040)*	.282 (.037)***	−.005 (.002)*		
Recognition of warning signals	.604 (.068)***	.932 (.063)***	.006 (.006)	−.018 (.009)*	−.027 (.008)***
Global risk perception	−.099 (.077)	−.008 (.071)	−.012 (.004)**		

	Predictor				
Outcome	**RTR**	**RTR+**	**Month**	**Month X RTR**	**Month X RTR+**
Global benefit perception	.081 (.065)	.009 (.060)	.004 (.003)		
Gist principles	.712 (.223)**	.288 (.208)	−.014 (.019)	−.064 (.028)*	−.060 (.026)*
Specific risk perception	−.015 (.044)	−.044 (.041)	.003 (.003)		
Quantitative risk perception	1.85 (1.16)	.648 (1.07)	−.053 (.064)		

Note. Cells contain parameter estimate (standard error). All estimates are adjusted for baseline score on the outcome, age, gender, ethnicity, and state where the curriculum was administered.

* $p<.05$,

** $p<.01$,

*** $p<0.001$.

Table 5

Adjusted Means of Outcome Variables by Intervention and Time of Assessment

	Time of Assessment			
	Postsurvey	**3 Months**	**6 Months**	**12 Months**
Behavior/behavioral intentions				
Number of sexual partners				
CONTROL	1.61 (.16)	1.83 (.16)	2.05 (.17)	2.49 (.19)
RTR	1.32 (.16)	1.54 (.16)	1.77 (.17)	2.21 (.19)
RTR+	1.26 (.16) [a]	1.48 (.16) [a]	1.70 (.17) [a]	2.15 (.19) [a]
Number of unprotected sex acts				
CONTROL	4.31 (.53)	6.33 (.61)	8.34 (.81)	12.37 (1.3)
RTR	4.02 (.52)	6.04 (.60)	8.05 (.81)	12.08 (1.3)
RTR+	3.85 (.52)	5.87 (.60)	7.88 (.80)	11.91 (1.3)
Prophylactic risk index				
CONTROL	.85 (.02)	.83 (.02)	.82 (.02)	.79 (.02)
RTR	.81 (.02)	.80 (.02)	.78 (.02)	.76 (.02)
RTR+	.84 (.02)	.83 (.02)	.81 (.02)	.78 (.02)
Intentions to have sex				
CONTROL	2.27 (.05)	2.30 (.05)	2.34 (.05)	2.40 (.08)
RTR	2.26 (.05)	2.32 (.05)	2.37 (.05)	2.48 (.08)
RTR+	2.17 (.05) [a]	2.24 (.05)	2.31 (.05)	2.45 (.07)
Intentions to use prophylaxis				
CONTROL	3.05 (.05)	3.06 (.05)	3.08 (.05)	3.10 (.06)
RTR	3.21 (.05)	3.22 (.05)	3.24 (.05)	3.27 (.06)
RTR+	3.17 (.05) [a]	3.16 (.05)	3.15 (.05)	3.12 (.06)
Attitudes and norms				
Attitudes toward sex				
CONTROL	1.65 (.04)	1.71 (.04)	1.76 (.05)	1.87 (.05)
RTR	1.68 (.04) [a]	1.73 (.04) [a]	1.78 (.04) [a]	1.89 (.05) [a]
RTR+	1.56 (.04) [a, b]	1.61 (.04) [a, b]	1.66 (.04) [a, b]	1.77 (.05) [a, b]
Attitudes toward prophylaxis				
CONTROL	2.95 (.03)	2.96 (.03)	2.98 (.03)	3.01 (.04)
RTR	3.10 (.03) [a]	3.10 (.03) [a]	3.11 (.03) [a]	3.11 (.04) [a]
RTR+	3.16 (.03) [a]	3.14 (.03) [a]	3.13 (.03) [a]	3.10 (.04)
Perceived sexual norms				
CONTROL	1.76 (.04)	1.82 (.04)	1.88 (.04)	2.00 (.04)
RTR	1.75 (.04)	1.81 (.03)	1.87 (.04)	1.99 (.04)
RTR+	1.68 (.04) [a, b]	1.74 (.03) [a, b]	1.80 (.04) [a, b]	1.92 (.04) [a, b]
Perceived parental sexual norms				
CONTROL	.95 (.05)	.99 (.05)	1.02 (.05)	1.09 (.05)
RTR	.95 (.05)	.99 (.05)	1.02 (.05)	1.09 (.05)

	Time of Assessment			
	Postsurvey	**3 Months**	**6 Months**	**12 Months**
RTR+	.94 (.05)	.96 (.05)	1.00 (.05)	1.06 (.05)
Perceived prophylactic norms				
CONTROL	3.02 (.04)	3.03 (.04)	3.03 (.04)	3.04 (.05)
RTR	3.11 (.04) [a]	3.12 (.04) [a]	3.12 (.04) [a]	3.14 (.05) [a]
RTR+	3.11 (.04) [a]	3.11 (.04) [a]	3.12 (.04) [a]	3.13 (.05) [a]
Motives				
Reasons to get pregnant				
CONTROL	1.67 (.03)	1.67 (.03)	1.67 (.03)	1.68 (.03)
RTR	1.62 (.03)	1.63 (.03)	1.63 (.03)	1.64 (.03)
RTR+	1.63 (.03)	1.64 (.03)	1.64 (.03)	1.65 (.03)
Reasons to have sex				
CONTROL	1.76 (.04)	1.76 (.04)	1.77 (.04)	1.77 (.05)
RTR	1.79 (.04)	1.79 (.04)	1.79 (.04)	1.80 (.05)
RTR+	1.73 (.04)	1.73 (.04)	1.74 (.04)	1.74 (.05)
Reasons not to have sex				
CONTROL	2.81 (.04)	2.77 (.04)	2.73 (.04)	2.65 (.04)
RTR	2.95 (.04) [a]	2.91 (.04) [a]	2.87 (.04) [a]	2.79 (.04) [a]
RTR+	2.97 (.04) [a]	2.93 (.04) [a]	2.89 (.04) [a]	2.81 (.04) [a]
Self-efficacy and behavioral control				
Self-efficacy (“saying no” to sex)				
CONTROL	2.83 (.05)	2.81 (.05)	2.79 (.05)	2.76 (.05)
RTR	3.02 (.05) [a]	3.00 (.05) [a]	2.98 (.05) [a]	2.95 (.05) [a]
RTR+	3.07 (.05) [a]	3.05 (.05) [a]	3.04 (.05) [a]	3.00 (.05) [a]
Prophylactic self-efficacy				
CONTROL	2.84 (.05)	2.85 (.04)	2.87 (.04)	2.89 (.05)
RTR	2.94 (.04) [a]	2.95 (.04) [a]	2.97 (.04) [a]	2.99 (.05) [a]
RTR+	3.02 (.04) [a]	3.03 (.04) [a]	3.04 (.04) [a]	3.07 (.05) [a]
Behavioral control (prophylaxis)				
CONTROL	2.81 (.04)	2.83 (.04)	2.85 (.04)	2.90 (.05)
RTR	2.94 (.04) [a]	2.96 (.04) [a]	2.99 (.04) [a]	3.03 (.05) [a]
RTR+	2.94 (.04) [a]	2.96 (.04) [a]	2.98 (.04) [a]	3.03 (.05) [a]
Gist/verbatim constructs				
Knowledge				
CONTROL	2.78 (.03)	2.80 (.03)	2.82 (.03)	2.87 (.04)
RTR	3.17 (.03) [a]	3.14 (.03) [a]	3.12 (.03) [a]	3.08 (.04) [a]
RTR+	3.27 (.03) [a, b]	3.25 (.03) [a, b]	3.22 (.03) [a, b]	3.17 (.04) [a]
Categorical risk perception				
CONTROL	2.88 (.04)	2.87 (.04)	2.85 (.04)	2.83 (.04)
RTR	2.99 (.04) [a]	2.97 (.04) [a]	2.96 (.04) [a]	2.93 (.04) [a]

	Time of Assessment			
	Postsurvey	**3 Months**	**6 Months**	**12 Months**
RTR+	3.16 (.04) [a, b]	3.15 (.04) [a, b]	3.14 (.04) [a, b]	3.11 (.04) [a, b]
Recognition of warning signals				
CONTROL	1.82 (.06)	1.83 (.06)	1.85 (.06)	1.89 (.08)
RTR	2.42 (.06) [a]	2.38 (.06) [a]	2.35 (.06) [a]	2.27 (.08) [a]
RTR+	2.75 (.06) [a, b]	2.68 (.06) [a, b]	2.62 (.06) [a, b]	2.49 (.07) [a, b]
Global risk perception				
CONTROL	1.98 (.07)	1.94 (.07)	1.91 (.07)	1.84 (.08)
RTR	1.88 (.07)	1.84 (.07)	1.81 (.07)	1.74 (.08)
RTR+	1.97 (.07)	1.94 (.07)	1.90 (.07)	1.83 (.08)
Global benefit perception				
CONTROL	1.50 (.06)	1.51 (.06)	1.52 (.06)	1.55 (.07)
RTR	1.58 (.06)	1.59 (.06)	1.60 (.06)	1.63 (.07)
RTR+	1.50 (.06)	1.52 (.06)	1.53 (.06)	1.56 (.07)
Gist principles				
CONTROL	10.61 (.20)	10.57 (.19)	10.53 (.20)	10.45 (.26)
RTR	11.32 (.20) [a]	11.09 (.19) [a]	10.86 (.20)	10.39 (.27)
RTR+	10.90 (.19) [a, b]	10.68 (.19) [a, b]	10.46 (.19)	10.02 (.25)
Specific risk perception				
CONTROL	.52 (.04)	.53 (.04)	.53 (.04)	.55 (.05)
RTR	.49 (.04)	.50 (.04)	.51 (.04)	.52 (.05)
RTR+	.53 (.04)	.54 (.04)	.55 (.04)	.57 (.05)
Quantitative risk perception				
CONTROL	5.85 (1.1)	5.70 (1.0)	5.54 (1.0)	5.22 (1.2)
RTR	7.70 (1.1)	7.54 (1.0)	7.38 (1.1)	7.07 (1.2)
RTR+	6.50 (1.1)	6.34 (1.0)	6.18 (1.0)	5.87 (1.0)

Note. Cells contain mean (standard error). All estimates are adjusted for baseline score on the outcome, age, gender, ethnicity, and state where the curriculum was administered.

[a] Significantly different from the control group at that time point.

[b] Significantly different from the RTR group at that time point.