Transtheoretical-based bullying prevention effectiveness trials in middle schools and high schools

Kerry E. Evers^a*, James O. Prochaska^b, Deborah F. Van Marter^a, Janet L. Johnson^a and Janice M. Prochaska^a

^aPro-Change Behavior Systems, Inc., Rhode Island, USA; ^bCancer Prevention Research Center, University of Rhode Island, USA

Background

Bullying threatens the physical and mental well-being of students across a broad range of schools. Tailored interventions based on the Transtheoretical model and delivered over the Internet were designed to reduce participation in each of three roles related to bullying (bully, victim and passive bystander).

Methods

Effectiveness trials were completed in 12 middle schools and 13 high schools in the USA. A diverse sample of 1237 middle (6th–8th grade, with 45.1% in the 7th grade; ages 11–14) and 1215 high school (9th–11th grade, with 41.6% in the 9th grade; ages 14–17) students were available for analyses.

Results

Analyses showed significant treatment effects for both intervention groups when compared to control for both the middle and high school programs.

Conclusions

Given the relative ease of dissemination, these programs could be applied as stand-alone practices or as part of more intensive interventions.

Keywords: Bullying prevention; Transtheoretical model; Stages of change; School based; Tailored communications; Internet

Background

Bullying is one of the most important behaviors threatening the physical and mental health of students across a broad range of grades and school systems (Nansel *et al.*, 2001, 2004). Bullying has been a major causal factor in tragic events, such as suicides by young students in Norway (Batsche & Knoff, 1994) and homicides by a growing number of students in the USA (Spivak & Prothrow-Smith, 2001). As a result, school

^{*}Corresponding author. Pro-Change Behavior Systems, PO Box 755, West Kingston, Rhode Island, RI 02892, USA. Email: kevers@prochange.com

systems have looked for evidence-based programs that could be implemented to reduce the rates and consequences of bullying.

The Olweus program from Norway is the most widely studied bullying prevention program. It involves highly intensive interventions that require considerable commitments of time and effort from teachers, counselors, principals, parents and students. While this program has produced some significant effects, the effects have been highly variable. In primary schools the effects have ranged from about 50% reduction in rates of bullies and victims in the first study in Bergin, Norway (Olweus, 1994) to 20% to 35% reduction in a more recent study (Olweus & Limber, 1999). In England (Whitney *et al.*, 1994) and Germany (Hanewinkel *et al.*, 1997) the reduction rates for bullies and victims was about 12% to 15%. In the USA there was a 25% reduction in rates of victims but a 5% increase in rates of bullies (Pepler *et al.*, 1994). In Belgium there were no significant reductions for either bullies or victims (Stevens *et al.*, 2000); and in Roagland, Norway, there was about a 4% increase in rates of bullies and victims (Roland, 1993).

In high school studies the results were more disappointing. In Belgium and Bergin, Norway, there were no significant effects on bullies or victims. In Roagland, Norway, there was about a 4% increase in rates of both bullies and victims. Only in England was there a significant decrease in bullies (12%) and victims (7%).

One of the major barriers to greater dissemination of the Olweus model program is the demand that it places on the staff and curriculum of schools. The prevention protocol included:

- 1. at the school level, a Bullying Prevention Coordinating Committee, an anonymous survey, a school conference day, improvement of supervision and outdoor environment, and meetings with parents;
- at the classroom level, classroom rules against bullying, positive and negative consequences, weekly classroom meetings and regular classroom meetings with parents;
- 3. at the student level, serious talks with bullies and victims;
- 4. at the family level, teacher and parent meetings. The programs were designed to be ongoing across multiple years and multiple grades (Olweus & Limber, 1999).

The project reported here was a series of effectiveness trials that applied Transtheoretical-based tailored programs that provided individualized and interactive computer interventions to populations of middle and high school students involved in bullying as bullies, victims and/or passive bystanders. The Transtheoretical model is a theory of behavior change that applies particular change processes like decision-making and reinforcement to help individuals progress at particular stages of change (Prochaska & DiClemente, 1983; Velicer *et al.*, 1985; Prochaska *et al.*, 1992; Prochaska & DiClemente, 1992). The intervention involved only three half-hour computer sessions during the school year for the students and a 10-page manual each for staff and parents that included optional activities. Nevertheless, based on population trials with other types of behaviors (Prochaska *et al.*, 1993, 2001a, b, 2004,

2005), this intervention was predicted to produce replicable results within a relatively narrow range of about 30% reduction for each of the three roles related to bullying.

Methods

Participants

Twelve middle schools (6th–8th grades) and 13 high schools (9th–12th grades) from across the USA were recruited from different types of communities (e.g. rural, working towns, suburbs, inner-city schools, urban blue collar and urban cores) with preference given to schools with higher percentages of students eligible for free lunches (about 45% of all students). Schools were matched on key variables (type of community, region of the country and percentage of students eligible for free lunches) and statistical procedures were used to control for any baseline differences. Power calculations were run to determine the sample size needed to have adequate statistical power and the analyses indicated that 140 students per treatment group at follow-up were needed to provide an 80% chance that a 10 percentage point difference between the treatment and control groups would be found significant at a probability (p-) level of 0.05.

Schools varied in the number of available classes with a range of three to 15 and a mean of nine classes. The demographics of the 1237 middle school students at baseline who were able to be matched at post-test show that there were more Hispanic students (27.4%) in the study than in the USA generally (about 14%).¹ There were fewer non-Hispanic Caucasians (40.5%) than in the USA and comparable percentages of African Americans (12%) and Asian Americans (4.4%). Approximately 48% of the students were eligible for free lunches and a little over 50% were female. The students were distributed across the 6th–8th grades (ages 11–14) with the largest percentage (45.1%) in the 7th grade.

The high school sample of 1215 students at baseline who were matched at posttest more clearly parallels the ethnic distribution in the USA. Approximately 42% were eligible for free lunches. About 55% were female and they were mostly from the 9th–11th grades (ages 14–17), with the 9th grade (41.6%) being the largest group.

Using self-report measures, students indicated participating in any of the three roles related to bullying (bully, victim, passive bystander), as well as their intention to stop. Based on behavior and intentions related to the three roles, participants indicated which of the five stages of change they were in. Only 4% of the middle school students were in the Maintenance stage (had not participated in any of the three roles related to bullying for six months or more), and 1% were in the Action stage (had not participated in any of the three roles related to bullying for less than six months). There were 16% in the Precontemplation stage (participating in at least one role related to bullying and not intending to stop in the next six months), about 27% were in the Contemplation stage (participating in at least one role related to bullying to stop in the next six months) and 52% were in the Preparation stage (participating in at least one role related to bullying to bullying but intending to stop in the next six months) and 52% were in the Preparation stage (participating in at least one role related to bullying but intending to stop in the next six months) and 52% were in the Preparation stage (participating in at least one role related to bullying but intending to stop in the next six months).

month). With the high school students, 6.5% were in Maintenance, 6% in Action, 25.6% in the Precontemplation, 25.8% in Contemplation and 41.5% in Preparation stages.

Design

A 3×2 experimental design crossed three experimental groups with two assessments. The three experimental groups were as follows:

- 1. *Control group*: this group received only pre- and post-tests and was composed of 483 middle school and 309 high school students.
- 2. *Treatment group 1*: this group, composed of 488 middle school and 375 high school students, received up to three Internet-based individualized and interactive sessions, a Staff Guide and a Family Guide, which are described above, as well as a post-test.
- 3. *Treatment group 2*: this group, composed of 266 middle school and 531 high school students, received the same protocol as group 1 and, in addition, received a pre-test prior to their first intervention. This group allowed for a replication of the interventions and a control for any effects due to a pre-test.

The pre-test and post-tests included the same assessment questions on role and stage of change only and were completed over the Internet. There were no significant differences in roles across the experimental groups at baseline. However, for the middle school sample there were significant differences on grade level ($\chi^2(8) = 46.11$, p < 0.001) and ethnicity ($\chi^2(6) = 123.78$, p < 0.001). In addition to grade level ($\chi^2(8) = 142.33$, p < 0.001) and ethnicity ($\chi^2(6) = 125.06$, p < 0.001), the high school sample also showed significant differences on gender ($\chi^2(2) = 9.13$, p < 0.05) and baseline stage of change ($\chi^2(8) = 26.41$, p < 0.01). The statistical analyses controlled for effects of all baseline measures.

Measures

All assessments were conducted by computer, which included both text and sound components. Outcome measures included the percentage of students who at post-test were no longer participating in each of the three roles related to bullying (bully, victim or passive bystander) and the percentage who were no longer participants when all three roles were combined. All three groups received these measures at pre-test or the start of the intervention for treatment group 1 and at post-test.

Self-report measures are the standard for assessing outcomes in the bullying prevention literature (Olweus & Limber, 1999). Studies have found that incidences of bullying reported to teachers, police or parents increase as a result of interventions, since victims and passive bystanders are encouraged to take effective action, such as reporting events to appropriate adults. When assessed anonymously or on computers where teachers do not know their responses, students have protection to report participation without interpersonal pressure from teachers, peers or parents. Such reporting is reflected in the fact that only about 10% of the middle school (MS) and high school (HS) students reported not participating in any role related to bullying.

Outcome measures

Role assessment. Three items were used to assess each role: (1) bully—treats people in mean ways and/or pushes, hits or kicks people; (2) victim—people treat them in mean ways and/or push, hit or kick them; and (3) passive bystanders—let people treat others in mean ways and let people push, hit or kick others while doing nothing to prevent or discourage such behaviors. At baseline in the middle schools there were 75.6% involved in roles as bullies, 80.6% as victims, and 64.4% as bystanders. In high schools there were 69.8% bullies, 71.2% victims and 74.7% bystanders. Obviously, many students participated in multiple roles.

Stage of change. Stage of change was determined by a single item matched to the roles students engaged in. Students were asked their intentions to make changes such that they would no longer be participating in specific roles. For example, individuals who engaged in bullying and passive bystander behavior were asked a single question on their intentions to stop each of these behaviors. If they had no intention to stop in the next six months, they were in the Precontemplation stage; if they intended to stop in the next month, they were in Contemplation; and if they intended to stop in the next month, they were in the Preparation stage. Individuals who indicated no participation in any role were asked how long they had been doing this and were placed in Action (less than six months) or Maintenance (more than six months). To be counted as an outcome success, students had to progress from participating in a role to being in Action or Maintenance at follow-up and no longer participating in that role.

In measurement development, the stages of change for participating in any of the roles of bullying were externally validated by their predicted relationship to key change variables—e.g. the pros and cons of changing and self-efficacy for not participating in bullying. Using standardized scores, the cons of changing were greater than the pros in Precontemplation, crossed over in Preparation and were higher in the Maintenance stage. Similarly, self-efficacy increased significantly across the stages of change (Evers *et al.*, 2003; Van Marter *et al.*, 2003).

Intervention measures

The two intervention groups were also assessed on the following measures that were relevant to their current stage of change at each intervention session. These assessments were used to tailor the intervention program to the needs of each student. The measures were developed in national pilot samples for middle school and high school separately. These studies also provided normative values and decision rules for generating feedback. These measures were used only to help guide students through the stages and were not used in outcome analyses.

402 K. E. Evers et al.

Pros and cons. Eight items each assessed the importance of advantages and disadvantages in an individual's decision to engage in each respective behavior. On a five-point Likert scale, importance ranged from 1 = N at all important to 5 = Extremely important. A sample pro item was: 'If I act respectfully, others will respect me back'; and a sample con item includes: 'Others will think I'm weak if I act respectfully'. These measures were developed in pilot studies specifically for middle school (pros $\alpha = 0.83$; cons $\alpha = 0.73$) and high school (pros $\alpha = 0.80$; cons $\alpha = 0.84$).

Processes of change. Ten processes of change for acting with respect towards self and others were assessed using 30 items that assessed cognitive, affective, experiential, and behavioral techniques used by individuals to facilitate the behavior change process. Students were asked to rate how often they used each technique during the previous month on a five-point Likert scale of frequency. Subscale internal consistency ranged from 0.65 to 0.80 for middle school and 0.61 to 0.77 for high school.

Self-efficacy. Temptations to participate in any of the three roles related to bullying were assessed for a variety of difficult situations on a five-point Likert scale (1 = Not at all tempted to 5 = Extremely tempted). In middle school, six items were used to assess two factors (negative affect and social situations) ($\alpha = 0.74$ and 0.78). For the high school population, eight items assessed the same two factors ($\alpha = 0.79$ and 0.82). Sample items included: 'How tempted would you be to act disrespectfully if you could get away with it?'.

Intervention materials

The Build Respect, Stop BullyingTM programs for middle school and high school are multi-component intervention packages. The primary component is a Transtheoretical model (TTM) tailored Internet-based expert system. Students were given the opportunity to interact with the program on three separate occasions. The computerized program was designed to be easy and engaging. The technical basis for this system relied on the integration of statistical, multimedia and database software. The system resided on a web server, and multimedia components were provided on a CD-ROM in order to minimize download time. All data collected by the program were encrypted and stored on a centralized database.

Students initiated the program by running the multimedia CD which brought the participant to the program website. The first time students interacted with the program (either at the pre-test or the first intervention), they were directed to register with the program by creating a login name based on personal information and a password. Once students registered for the program, logged in and consented to be part of the research, they were given instructions on how the program worked. The program led the student through a series of screens that included assessment questions, feedback on their answers, images and Quick Time movies that were all tailored to the specific needs of the student.

The program began with an assessment of behaviors and roles the students played in bullying. As participants answered the questions, the system scored their assessments, referred to the decision rules and database and provided immediate onscreen feedback. To generate individualized expert system feedback, students were then assessed on all TTM constructs relevant to their specific stage of change at each session. The assessment provided the input for the expert system and was the basis for the tailoring and individualization of the feedback the student received. For each stage of change, the program analyzed the participant's responses to the particular constructs for the stage and determined where they were in relation to a predetermined level of construct use to optimize movement to the next stage. The expert system then produced individualized feedback that was provided to the participant through text and graphical feedback on their computer screen.

The first intervention time point provided students with normative feedback only. This normative feedback compared the individual's use of change principles and processes to peers who were most successful in progressing. Sessions 2 and 3 provided both normative (compared to peers who progressed the most) feedback on their current use of TTM variables and ipsative (compared to self) feedback on how they were progressing since the last interaction with the program. For example, the ipsative feedback might inform particular students they had progressed two stages since their last interaction, which meant they had about tripled their chances that they would be free from any bullying role in the next six months. Feedback also reinforced them for any progress they were making on any change variable, and what behavioral strategies they could emphasize to progress to the next stage. The program was retailored to the individual student's needs at each session.

The program included text and multimedia components. The text was also read to the students who were advised to only choose the sound on version if they had headphones available. Images on the screen were also matched to the specific feedback that was being provided. Finally, short movies of students giving testimonials about bullying or changing were provided at specific times throughout the program.

A 10-page Family Guide was mailed to participants' families and provided brief information about the program and general information on bullying. In addition, school staff received a 10-page Staff Guide, which included general information on bullying and how to support student change, classroom activities and information on how to work with parents. They were also provided access to a website that included activities they could incorporate into the classroom. Teachers were not provided training and their main responsibilities were to assist students in starting and completing the Internet program.

Procedure

Schools were recruited using proactive outreach techniques, as follows. Once a school agreed and was assigned to an experimental group, they sent home introductory packets, which included consent forms, an introduction letter and a Family Guide for those schools in the intervention groups. Passive consent was used for all but three middle schools and forms were available to guardians to send back if they did not want a student to participate. All procedures were reviewed and approved by an

Institutional Review Board. The control group and treatment group 2 were given a pre-test when they were ready to participate. Treatment groups 1 and 2 were given the first Internet session as soon as they and the program were ready for the study. The treatment groups were told that their remaining two sessions should be completed at approximately three-month intervals. The reality was that the sessions were completed within a typical two-month window ranging from one to three months.

Teachers were provided the multimedia CD and the Staff Guide. They had access to a website to download worksheets to be used in the classroom, if they chose to. The CD brought the user to the program website and the program was delivered over the Internet. The multimedia CD allowed for fewer delays due to downloading individual multimedia components.

Schools and teachers had considerable variability in their technical capabilities for delivering Internet-based programs. Problems included classes not being able to obtain adequate computer lab time; teachers using the wrong URL to access the program; schools having firewalls and security settings that limited program use; schools having very slow Internet connections that limited the program use; and students not recording their login codes and having to change codes. These problems have since been corrected and the specifications are made very clear concerning the technical requirements needed to run the program.

Matching/retention

Problems in accessing the program produced unique challenges in matching data records and in determining retention rates. The use of login codes to protect confidentiality produced a research challenge that we had not faced in any of our previous population trials. Students who forgot their confidential login codes could continue in the study by entering a new code at a subsequent session. Because they used a new code, they were treated as a new participant and started on the first session. If these students forgot their new codes at the third session, they could continue with a third new code. Problems emerged when we matched their follow-up assessment with their first session assessment. It would appear that two-thirds of such students were not retained in the study, when the problem was that we could only match their follow-up assessment to one of their three 'first'-session assessments. Given this problem we could not calculate traditional retention rates. We could calculate the rates of matched records-that is, those whose login codes for their 'first' session assessment matched their codes for their post-test assessment session. The unmatched records included students who used new login codes and students who were not retained in the study. The problems with matched and unmatched records were taken into account in our plans for analyses.

For the middle school study all three groups had significantly different rates of matching ($\chi^2(2) = 225.24$, p < 0.001). The control group (59.2%) showed the highest rate followed by treatment group 1 (48.8%) and treatment group 2 (25.8%). With the high school study, the control group (50.3%) and treatment group 1 (50.5%) had about equal rates of matched records, while treatment group 2 (34.6%) was significantly lower ($\chi^2(2) = 65.17$, p < 0.001).

Analysis plan

Due to the differential ability to match by treatment group and other methodological complexities, we relied on multiple analyses of the data. The first set involves a complete case analysis on students who were able to be matched, for middle school (n = 1230) and high school (n = 1203). Differences in proportions at post-test and random effects logistic models were performed. The random effects models controlled for demographic differences at baseline and the school being the unit of assignment.

The second set of analyses used complete observations. Intent-to-treat analyses examined effects in all records rather than only those who were able to be matched, for middle school (n=2839) and high school (n=2713). The last observation carried forward (LOCF) method of replacement was used where a missing post-test value was replaced with the last known observation for each record. The LOCF method was appropriate for the design of these studies; the approach is conservative in that it assumes no change over time when the hypothesis of interest tests that people will improve over time, so it will predictably produce lower estimates of effect sizes (Steiner, 2002).

Results

Middle school results

Complete case analyses: difference in proportions

Results for all three roles combined. Figure 1 presents the percentage of the three groups in the middle school study who reported being in Pre-action at baseline, who progressed to Action or Maintenance for all three roles combined at post-test. Approximately 22% of treatment group 2 and 18% of treatment group 1 had progressed to Action or Maintenance compared to 5% of the control group. Both intervention groups showed significantly higher proportions no longer participating in any of the three roles (moved to Action or Maintenance) than the



Figure 1. Percentages of middle school students in Action or Maintenance stages for all three roles combined

control group (treatment 1 versus control: Z = 6.45, p < 0.001, h = 0.42; treatment 2 versus control: Z = 6.38, p < 0.001, h = 0.50). The *h*-statistic is an estimate of effect size and is comparable to Cohen's *d* (Cohen, 1988). The odds ratio for the two intervention groups combined compared to the control group was 4.38, meaning that the intervention groups were more than four times as likely as the control group to progress to Action or Maintenance and not participate in any role related to bullying.

Results for individual roles (bully, victim, passive bystander). Figure 2 presents percentages of all three groups, who reported being a bully at baseline and progressed to not being a bully at the post-test. Approximately 32% of treatment group 2 and 28% of treatment group 1 had progressed to not being a bully at post-test, compared to 19% of the control group. Both intervention groups showed significantly higher proportions of individuals becoming non-bullies than the control group (treatment 1 versus control: Z=2.88, p < 0.01, h=0.21; treatment 2 versus control: Z=3.47, p < 0.001, h=0.30).

For students who reported being a victim at baseline, approximately 35% of treatment group 2 and 30% of treatment group 1 had progressed to not being a victim at post-test, compared to 17% of the control group. Both intervention groups showed significantly higher proportions no longer reporting being victims than the control group (treatment 1 versus control: Z=4.33, p < 0.001, h=0.31; treatment 2 versus control: Z=4.89, p < 0.001, h=0.41). For students who reported being a passive bystander at baseline, approximately 34% of treatment groups 1 and 2 had progressed to taking appropriate action to prevent bullying at post-test, compared to about 21% of the control group. Both intervention groups showed significantly higher proportions no longer being bystanders than the control group (treatment 1 versus control: Z=3.59, p < 0.001, h=0.29; treatment 2 versus control: Z=2.98, p < 0.01, h=0.28).

Complete case analyses: random effects logistic models

Results for all three roles combined. A random effects logistic model on the proportion of people in Action or Maintenance was run using the baseline demographics and stage



Figure 2. Percentages of middle school students reporting no bullying behavior

of change as covariates. The intraclass correlation for the school level after controlling for the covariates was 0.018. The intraclass correlation effects were also controlled for in this analysis. At pre-test treatment group 2 was equivalent to the control group (t(2622) = -0.25, p > 0.01). At the post-test, both intervention groups showed significantly higher proportions of people in Action/Maintenance than the control group (treatment 1: t(2622) = -6.33, p < 0.001; treatment 2: t(2622) = -7.55, p < 0.001). By the second intervention continuing on through the post-test, there were no significant differences between the two intervention groups.

Results for individual roles (bully, victim, passive bystander). Similar random effects logistic models were run for each of the individual roles. The intraclass correlation for school level after controlling for the covariates was 0.036 for bully, 0.018 for victim and 0.014 for bystander. Again, the analyses controlled for intraclass correlation effects. For all three roles, treatment group 2 was equivalent to the control group at pre-test. However, at post-test treatment group 2 had significantly fewer people reporting being a bully, victim or bystander. Treatment group 1 showed similar significant differences from the control group at post-test for all three roles. There were no significant differences between the two treatment groups at any of the intervention time points or post-test for any of the roles.

Complete observations: intent-to-treat analysis

Intent-to-treat analyses were used with all observations and the last observation carried forward (LOCF) method of replacement was used where there was a missing post-test variable. Both intervention groups show significantly more people in Action/ Maintenance than the control group. Similarly, both intervention groups were significantly higher than the control group for the proportion of people who no longer reported being a bully, a victim or a bystander.¹

High school results

Complete case analyses: difference in proportions

Results for all three roles combined. Figure 3 presents the results of all three groups in the high school study who reported being in pre-action at baseline who progressed to Action or Maintenance for all three roles combined at post-test. Approximately 29% of both treatment groups had progressed to Action or Maintenance and were not participating in any of the roles related to bullying, compared to 10% of the control group. Using arcsine transformation comparisons between proportions both intervention groups showed significantly higher proportions moving to Action or Maintenance than the control group (treatment 1 versus control: Z = 6.17, p < 0.001, h = 0.49; treatment 2 versus control: Z = 6.86, p < 0.001, h = 0.51). The odds ratio for the two intervention groups combined compared to the control group was 3.89, meaning that the intervention groups were almost four times as likely to progress to Action or Maintenance than was the control group.

Results for individual roles (bully, victim, passive bystander). Figure 4 presents the percentage of all three groups who reported being a bully at baseline who progressed to not being a bully at post-test. Approximately 42% of treatment group 1 and 38% of treatment group 2 progressed at post-test compared to 21% of the control group. Both intervention groups showed significantly higher proportions of individuals becoming non-bullies than the control group (treatment 1 versus control: Z=5.06, p < 0.001, h=0.46; treatment 2 versus control: Z=4.33, p < 0.001, h=0.37).

Of students who reported being a victim at baseline, approximately 40% of treatment group 1 and 37% of treatment group 2 had progressed to not being a victim at post-test compared to 22% of the control group. Again, both intervention groups showed significantly higher proportions no longer reporting being victims than the control group (treatment 1 versus control: Z = 4.36, p < 0.001, h = 0.39; treatment 2 versus control: Z = 3.93, p < 0.001, h = 0.33). Of the students who reported being a passive bystander at baseline, approximately 43% of treatment group 1 and 41% of treatment group 2 had progressed to taking appropriate action to prevent bullying at



Figure 3. Percentages of high school students in Action or Maintenance stages for all three roles combined



Figure 4. Percentages of middle school students reporting no bullying behavior

post-test, compared to 24% of the control group. Both intervention groups showed significantly higher proportions no longer being bystanders than the control group (treatment 1 versus control: Z=4.91, p < 0.001, h=0.43; treatment 2 versus control: Z=4.66, p < 0.001, h=0.38).

Complete case analyses: random effects logistic models

Results for all three roles combined. A random effects logistic model was run for the high school sample using the baseline demographics as covariates. The intraclass correlation for the school level after controlling for the covariates was 0.023, which was also controlled in this analysis. At the pre-test, treatment group 2 was equivalent to the control group (t(2949) = -0.08, p > 0.01). At the post-test, both intervention groups showed significantly higher proportions of people in Action/Maintenance than the control group (treatment 1: t(2949) = -5.33, p < 0.001; treatment 2: t(2949) = -5.69, p < 0.001). By the third intervention continuing on through the post-test, there were no significant differences between the two intervention groups.

Results for individual roles (bully, victim, passive bystander). Similar models were run for each of the three roles—bully, victim and bystander. The intraclass correlation for school level after controlling for the covariates was 0.017 for bully, 0.020 for victim and 0.0 for bystander. For all three roles, treatment group 2 was equivalent to the control group at pre-test. However, at post-test treatment group 2 had significantly fewer people reporting being a bully, victim, or bystander than the control group. Treatment group 1 showed similar significant differences from the control group at post-test for all three roles. The two intervention groups differed at the first intervention for bully and the second intervention for victim. Otherwise, the treatment groups were equivalent across time points for the three roles.

Complete observations: intent-to-treat analyses

Again, the intent-to-treat analyses used all observations and the last observation carried forward (LOCF) method of replacement for missing post-test values. Both intervention groups showed significantly more people in Action/Maintenance than the control group. Similarly, both intervention groups were significantly higher than the control group for the proportion of people who no longer reported being a bully, a victim or a bystander.¹

Discussion

An extensive and sophisticated series of analyses produced a consistent pattern of results that indicated that the interventions produced significant reductions in the percentage of students in middle schools and high schools who participated in bullying. The same pattern occurred for the roles of bully, victim and passive bystander. Taken together, these results indicate that individualized and interactive TTM-based interventions of relatively low intensity can be effective in reducing

self-reported rates of middle school and high school students involved in bullying as bullies, victims and/or passive bystanders.

The results are particularly important given the relatively low intensity of only three half-hour sessions and the relatively low demands the program places on school personnel who want to implement bullying prevention programs. The greater the intensity, and the greater the demands on school staff for implementation, the greater the barriers there are for widespread dissemination of such prevention programs. The results were especially striking for high school populations, since the evidence for the effectiveness of the most studied and disseminated program has been mostly negative.

Confidence in these results is enhanced by the amount of replication built into this research. First, there were replications of significant reductions across the middle school and high school studies. Next, there were replications within each study not only in terms of significant reductions, but also the magnitude of the reductions. Both treatment groups within both studies produced rates of reduction that were always within seven percentage points of each other, whether complete case or intent-to-treat analyses were used. Furthermore, there were replications across three different roles related to bullying and across all three roles combined. Being able to predict not only significant reductions, but also the relative magnitude of the reductions, is important for the science and practice of bullying prevention. Since the Olweus program is as good and intensive as it is, the magnitude of significant reductions ranged from zero to 50%. We can only speculate as to why there is such range in reductions. It is not due to the size of the samples studied since most of the studies on the Olweus program were larger than the present project. The larger the sample size, the more stable the estimates should be for expected results. One factor could be that the studies were done in different countries, but there were still fairly wide ranges of results within England and within Norway. Also, there was little difference in reductions across Germany and England for elementary school students.

Variability in the magnitude of results could be related to the strengths and weaknesses of school-based prevention programs that are implemented by teachers. Strengths include the personal relationships that teachers can have with students and their parents and the teachers' abilities to respond to personal concerns of particular students or parents. Weaknesses include the variability in time that teachers have available for training, supervision and implementation of more intensive programs. Such variability in time and resources available for implementation can lead to considerable variability in outcomes (Botvin *et al.*, 1995; Peterson *et al.*, 2000).

With computer-based individualized and interactive interventions, there are no variations in the quality of the expert guidance that is provided. Quality is built into the computer. There is huge variability in the specific feedback and guidance that is provided to each participant, because the programs are tailored to the assessment of each individual at each session. Although there are only three sessions, each one is tailored to the needs of each student rather than relying on one-size-fits-all programs designed for the entire class. Furthermore, the guidance given to each student is not based on the clinical judgment of a clinician or teacher, since the quality of such judgments can be highly variable. Instead, the normative and ipsative feedback

provided is based on statistical decision-making that is known to outperform clinical judgments.

In this effectiveness trial, there was variability in the duration between sessions and the number of sessions delivered. Such variability produces error or noise that would be expected to reduce treatment effects. The fact that there were significant treatment effects in spite of such variability is important for dissemination. As highlighted by Glasgow *et al.* (2003), significant results in effectiveness trials indicate that the interventions are likely to be robust enough to generalize to variable conditions in the real world outside a highly controlled research setting.

How were significant treatment effects found in spite of significant variability due to baseline demographics and stage of change, school-level intra-class correlations and differential rates of complete cases? To find significant results after controlling for such variability generally requires robust treatment effects. The odds ratios of 4.38 and 3.89 are the highest we have found for TTM-tailored interventions for any behavior in any population. These highest odds ratios hold when comparing across studies that used complete case analyses, self-reports and students in middle school or high school (Prochaska *et al.*, 1993, 2001a, b, 2004, 2005; Jones *et al.*, 2003).

Limitations

One limitation is the representativeness of the samples. The ethnic and racial demographics of the high school sample were similar to those of the USA generally, and the middle school sample had higher than average representation of Hispanic students. Nevertheless, the samples were not selected to be representative of schools and students across the USA. This limitation raises concerns about how broadly the results can be generalized. These concerns are tempered by the amount of replication built into the project. Some sciences, like psychology, rely on replications as the basis for generalization more than representativeness.

The time of the post-test can be seen as a limitation, with the average post-test occurring about one month after the last interaction. Unfortunately, other research on bullying prevention programs does not make it clear whether shorter duration of evaluations helps or hurts a study. In two studies that did evaluations at six and seven months, three of the six assessments were not significant (Melton *et al.*, 1998; Olweus & Limber, 1999). In those studies that did evaluations at 18 months to three years, nine assessments were significant, and 11 were not significant (Roland, 1993; Pepler *et al.*, 1994; Olweus & Limber, 1999; Stevens *et al.*, 2000). It should be noted that in all those studies the intervention was ongoing, and none of the studies did post-test evaluations after the programs had finished.

The follow-up limitation is tempered somewhat by the pattern of results over time with TTM-tailored interventions. In population-based effectiveness trials on behaviors like smoking, diet and sun exposure, the most common pattern over time was that differences between treatment and control groups increased over 12 to 18 months after the treatment ended (Prochaska *et al.*, 1993, 2001a, b, 2004, 2005). The other pattern was that significant treatment effects, as with stress management, were

sustained 12 months after treatment ended (Evers *et al.*, submitted). In all trials significant effects at the end of treatment were still found at 12–18-month follow-ups.

However, there are no data that have assessed the sustainability of treatment effects of bullying programs long after the treatment has ended. This limitation is also tempered by some value considerations. The issue of sustainability of treatment effects is especially important when the targeted health behavior produces mostly long-term risks such as smoking, diet or sun exposure. With bullying, however, the risks and negative consequences are immediate and programs can be valuable if they produce significant reductions in bullying just while they are in operation. Furthermore, given the value of such programs, there is no inherent reason why they have to be discontinued.

Another limitation is reliance on self-report measures for each of the three roles related to bullying. This is a limitation of most research on bullying prevention, in part because there is no established gold standard for objective measurement. Reports to schools of bullying incidents are objective measures, but such reports are likely to increase in intervention groups as victims and bystanders become more empowered to take such actions. Nevertheless, there are concerns about self-reports being sensitive to experimenter demands or pressures from teachers or counselors to report more positive changes. These concerns are tempered somewhat by the fact that the programs were implemented by computers rather than teachers or counselors. Further, the programs were stage-based and designed to respect whatever stage participants were in and not to pressure them to take action. Rather, the program communicated that the process would work best if participants took action whenever they were ready rather than in response to outside pressures. Still, there are limitations to relying just on self-reported outcomes.

Future directions

Given replicated results with a low intensity intervention that has produced significant and sizeable reductions in all three roles related to bullying, a key question is how these interventions can be enhanced. Would continuing these programs through a second school year continue to produce greater reductions in the percentages of students involved in bullying as bullies, victims and passive bystanders? Would combining these programs with more intensive interventions, like the esteemed Olweus program, produce greater reductions in each of the three roles related to bullying? Would the combination of teacher-led classroom activities providing more personalized approaches and computer-delivered programs providing more individualized guidance produce greater impacts on bullying?

As the bullying prevention program is disseminated, we will continue to evaluate its impacts in real-world applications. Because these are Internet-based programs, we will be able to evaluate the progress, or lack of progress, that particular populations and subpopulations make as they participate in the program. We will seek to identify conditions under which it is not performing up to predicted impacts and seek to correct those conditions. We will also seek to identify conditions under which it is performing beyond expectations and work to see if we can replicate those conditions to produce breakthroughs in the practice of bullying prevention.

Note

1. Tables for middle and high school students are available from the first author at kevers@prochange.com for demographics and each of the statistical analyses.

References

- Batsche, G. M. & Knoff, H. M. (1994) Bullies and their victims: understanding a pervasive problem in the schools, *School Psychology Review*, 23(2), 165–174.
- Botvin, G. J., Baker, E., Dunesbury, L., Botvin, E. M. & Diaz, T. (1995) Long-term follow-up results of a randomized drug abuse prevention trial in a white middle-class population, *Journal of American Medical Association*, 273, 1106–1112.
- Cohen, J. (1988) Statistical power analysis for the behavioral sciences (2nd edn) (Hillsdale, NJ, Lawrence Erlbaum).
- Evers, K. E., Prochaska, J. O., Mauriello, L. M., Johnson, J. L., Padula, J. A. & Prochaska, J. M. (submitted) A randomized clinical trial of a population-based stress management intervention: 18-month outcomes.
- Evers, K. E., VanMarter, D. F., Johnson, J. L., Eastwood, A. L., Prochaska, J. M. & Prochaska, J. O. (2003) Development and validation of transtheoretical model variables applied to bullying prevention among high school students, paper presented at the 24th Annual Scientific Sessions of the Society of Behavioral Medicine, Salt Lake City, Utah, March.
- Glasgow, R. E., Lichtenstein, E. & Marcus, A. C. (2003) Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition, *American Journal of Public Health*, 93, 1261–1267.
- Hanewinkel, R. & Knaach, R. (1997) Mobbing: Gewaltprövention in the schule in Schleswig-Holstein: report (Kiel, Landesenstitut Schleswig-Holstein für Praxis und Theorie der Schule).
- Jones, H., Edwards, L., Vallis, T. M. et al. (2003) Changes in diabetes self-care behaviors make a difference in glycemic control: the Diabetes Stages of Change (DiSC) study, *Diabetes Care*, 26, 732–737.
- Melton, G. B., Limber, S. P., Cunningham, P. et al. (1998) Violence among rural youth: final report to the Office of Juvenile Justice and Delinquency Prevention (Washington, DC, US Department of Justice).
- Nansel, T. R., Craig, W., Overpeck, M., Saluga, G., Ruan, W. J. & Health Behaviour in Schoolaged Children Bullying Analyses Working Group (2004) Cross-national consistency in the relationship between bullying behaviors and psychosocial adjustment, *Archives of Pediatrics and Adolescent Medicine*, 158(8), 730–736.
- Nansel, T. R., Overpeck, M., Pilla, R. S., Ruan, W. J., Simons-Morton, B. & Scheidt, P. (2001) Bullying behaviors among US youth, *Journal of the American Medical Association*, 285(16), 2094–2100.
- Olweus, D. (1994) Bullying at school: long-term outcomes for the victims and an effective schoolbased intervention program, in: R. Huesmann (Ed.) *Aggressive behavior: current perspectives* (New York, Plenum Press).
- Olweus, D. & Limber, S. (1999) *Blueprints for violence prevention: 9, Bullying prevention programme* (Boulder, CO, Center for the Study and Prevention of Violence).
- Pepler, D. J., Craig, W. M., Ziegler, S. & Charach, A. (1994) An evaluation of an anti-bullying intervention in Toronto schools, *Canadian Journal of Community Mental Health*, 13, 995–110.
- Peterson, A. V., Kealey, K. A., Mann, S. L., Marek, P. M. & Sarason, I. G. (2000) Hutchinson smoking prevention project: long-term randomized trial in school-based tobacco use prevention results on smoking, *Journal of the National Cancer Institute*, 92, 1979–1990.

- Prochaska, J. O. & DiClemente, C. C. (1983) Stages and processes of self-change of smoking: toward an integrative model of change, *Journal of Consulting and Clinical Psychology*, 51, 390– 395.
- Prochaska, J. O. & DiClemente, C. C. (1992) Stages of change in the modification of problem behaviors, in: M. Herson, R. Eisler & P. Miller (Eds) *Progress in behavior modification* (Sycamore, IL, Sycamore).
- Prochaska, J. O., DiClemente, C. C. & Norcross, J. C. (1992) In search of how people change: applications to addictive behaviors, *American Psychologist*, 47(9), 1102–1114.
- Prochaska, J. O., DiClemente, C. C., Velicer, W. F. & Rossi, J. S. (1993) Standardized, individualized, interactive and personalized self-help programs for smoking cessation, *Health Psychology*, 12, 399–405.
- Prochaska, J. O., Velicer, W. F., Fava, J. L. et al. (2001a) Counselor and stimulus control enhancements of a stage-matched expert system intervention for smokers in a managed care setting, *Preventive Medicine*, 32, 23–32.
- Prochaska, J. O., Velicer, W. F., Fava, J. L., Rossi, J. S. & Tsoh, J. Y. (2001b) Evaluating a population-based recruitment approach and a stage-based expert system intervention for smoking cessation, *Addictive Behaviors*, 26, 583–602.
- Prochaska, J. O., Velicer, W. F., Redding, C. A., Rossi, J. S., Goldstein, M., DePue, J., Greene, G. W., Rossi, S. R. & Sun, X. (2005) Stage-based expert systems to guide a population of primary care patients to quit smoking, eat healthier, prevent skin cancer and receive regular mammograms, *Preventive Medicine*, 41, 406–416.
- Prochaska, J. O., Velicer, W. F., Rossi, J. S. *et al.* (2004) Impact of simultaneous stage-matched expert systems for multiple behaviors in a population of parents, *Health Psychology*, 23, 503– 516.
- Roland, E. (1993) Bullying: a developing tradition of research and management, in: D. P. Tattum (Ed.) Bullying in schools (Oxford, Heinemann Educational).
- Spivak, H. & Prothrow-Smith, D. (2001) The need to address bullying—an important component of violence prevention, *Journal of the American Medical Association*, 285, 2131–2132.
- Steiner, D. L. (2002) The case of the missing data: methods of dealing with dropouts and other research vagaries, *Canadian Journal of Psychiatry*, 47, 68–75.
- Stevens, V., DeBourdeauhuij, I. & Van Oost, P. (2000) Bullying in Flemish schools: an evaluation of anti-bullying intervention in primary and secondary schools, *British Journal of Educational Psychology*, 70, 195–210.
- Van Marter, D. F., Evers, K. E., Johnson, J. L., Eastwood, A. L., Prochaska, J. M. & Prochaska, J. O. (2003) Measure development of Transtheoretical model constructs for bullying prevention in middle schools, paper presented at the 24th Annual Scientific Sessions of the Society of Behavioral Medicine, Salt Lake City, Utah, March.
- Velicer, W. F., DiClemente, C. C., Prochaska, J. O. & Brandenburg, N. (1985) Decisional balance measure for assessing and predicting smoking status, *Journal of Personality and Social Psychology*, 48, 1279–1289.
- Whitney, L., Revers, I., Smith, P. & Sharp, S. (1994) The Sheffield project: methodology and findings, in: P. Smith & S. Sharp (Eds) School bullying: insights and perspectives (London, Routledge).

Copyright of Educational Research is the property of Routledge and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use. Copyright of Educational Research is the property of Routledge and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.