Can public campaigns effectively change psychological determinants of safer sex? An evaluation of three Dutch campaigns

M. C. Yzer, F. W. Siero and B. P. Buunk

Abstract

This study evaluated the 1994, 1995 and 1996 Dutch safer sex campaigns as to their effectiveness in terms of improved attitudes, perceived social norms, self-efficacy and intentions regarding safer sex. The hypotheses were tested that variables become more positive when campaigns are conducted and less positive when campaigns are discontinued. A comprehensive design, including a baseline–post-test/post-test-only group design and a longitudinal or multiple assessment group design, was employed to exclude testing effects, history effects, cultural changes and sample differences as alternative explanations. The results show that despite high baseline levels, the campaigns positively affected all variables. Importantly, levels of all variables decreased when no campaign was conducted. It is concluded that campaigns are needed to maintain high levels of determinants of safer sex and that future campaign goals should be formulated in terms of stabilization instead of growth.

Introduction

Since the first use of mass media campaigns for health promotion, there has been considerable debate about the effectiveness of such campaigns in changing health-related behaviors in the population. In an effort to enhance their effectiveness, campaigns have been increasingly based on various models of health behavior [cf. (Fisher and Fisher, 1992)]. In the context of safer sex campaigns, Ajzen’s Theory of Planned Behavior (TPB) (Ajzen, 1991) is one of the most widely used models. The model postulates that the intention to engage in safer sex is a direct predictor of actual safer sex behavior in the future. In turn, the intention is assumed to be based upon attitudes, social norms and self-efficacy with respect to safer sex. Attitudes pertain to the evaluation of safer sex in terms of good versus bad, social norms concern estimates of the extent to which the reference group approves of safer sex and self-efficacy pertains to perceptions of the ability to actually carry out the behavior. Whereas Ajzen originally labeled the latter variable as personal control, this variable is highly similar to Bandura’s (Bandura, 1994) concept of self-efficacy. Because self-efficacy has become a more widely used label, in particular in the context of health-related behavior, in this paper we prefer to use this term rather than personal control. Numerous studies have examined the ability of the TPB to predict safer sex intentions and behaviors [e.g. (Basen-Enquist and Parcel, 1992; Fishbein et al., 1992; Walter et al., 1994; Fisher et al., 1995; Morrison et al., 1995)]. In the present research, we examine to what extent attitudes, social norms, self-efficacy and intentions with respect to safer sex became more positive in response to the 1994, 1995 and 1996 Dutch safer sex campaigns. For a comprehensive evaluation of the campaigns, we also consider whether the same set of target variables became less positive in the period in which no campaign was conducted.
General features of the 1994, 1995 and 1996 Dutch public safer sex campaigns

The 1994, 1995 and 1996 safer sex campaigns were part of a prevention program entitled ‘I have Safe Sex or No Sex’. In this program, campaigns were conducted aimed at the general public, small-scale prevention projects were conducted aimed at specific target groups, and basic information about AIDS and other STDs was made available throughout the year. The public campaigns were conducted during each summer. In this article, we focus solely on the effects of the campaigns and not on the effects of other prevention activities. General features of the campaigns were the recommendation of condom use as the most important means to ensure safer sex, equal attention to discussing condom use with partners and to actual condom use, the integration of AIDS prevention with STD prevention, and agenda setting as the major overall objective. In a campaign context, agenda setting refers to the process of keeping safer sex issues in the public awareness (Roberts and Maccoby, 1985). As the Dutch safer sex campaigns had as explicit goals to influence attitudes, social norms, self-efficacy and intentions with respect to safer sex, the present evaluation study focused on changes in this set of target variables. Moreover, because the campaigns paid explicit attention to both actual condom use and discussing condom use with new partners, we used measures for the target variables that referred to both types of behaviors.

Evaluation of public safer sex campaigns

Research assessing the effects of campaigns aimed at fostering safer sex in the population is particularly important for appropriately designing future campaigns (Winett et al., 1990). Nevertheless, there is only a small number of studies that have evaluated the effects of such campaigns. For example, out of the 6000 papers presented at the 1998 World AIDS Conference, less than 20 could be characterized as public campaign evaluations. The paucity of published quantitative evaluation studies is partly due to the methodological challenges that evaluation researchers are facing. Indeed, because of the very nature of a public campaign, it is impossible to isolate a true control group that can be compared to an experimental group. Research examining the effects of interventions aimed at fostering safer sex that did include a control group has been limited to relatively small-scale interventions [e.g. (Van den Eijnden et al., 1998; Yzer et al., 1998)]. In studies on the effects of public campaigns, the design that most closely approached an experimental design was one in which an intervention was implemented in one geographic area (experimental group), while another area served as control group [e.g. (Borgia et al., 1997; Fishbein et al., 1997)]. It is evident that in these studies participants were not randomly assigned to an experimental or control group, which threatens the validity of conclusions about the effects of the intervention under consideration.

Such randomization problems were less of an issue in the design that Middlestadt et al. (Middlestadt et al., 1995) employed in their evaluation of a campaign aimed at fostering safer sex that was conducted in the Caribbean. Because the campaign was conducted nation-wide, it was not possible to differentiate between an experimental and a control group on the basis of geographic areas in which the campaign was implemented or was not implemented. Instead, Middlestadt et al. measured the extent to which respondents were exposed to the campaign. Because exposure to the campaign appeared not to be dependent on relevant confounding variables, such as age and gender, relatively weakly exposed persons could be compared to relatively strongly exposed persons in order to investigate campaign effects. Differences between strongly versus weakly exposed individuals were interpreted as campaign effects. Although the use of degree of exposure to examine campaign effects has definite advantages, the cross-sectional character of the Middlestadt et al. study renders it impossible to make inferences about the development of the effects of the campaign over time. Such inferences about the development of the effects of a campaign require assessments of the same variables among the same individuals before and after the campaign. In most studies that have reported longitudinal data on the effects of a
Public safer sex campaigns

public campaign, new samples were recruited for each of the baseline tests and post-tests [e.g. De Vroome et al., 1994]). Therefore, in these studies it cannot be unequivocally determined whether results reflect campaign effects or differences between samples.

Increasing the validity of causal inferences about public campaign effects

Given the many problems associated with making valid inferences about the effects of campaigns aimed at fostering safer sex, the present research aimed to exclude a number of alternative explanations to convincingly demonstrate that results are indeed effects of campaigns. In their authoritative work on quasi-experimental research designs, Cook and Campbell (Cook and Campbell, 1979) outlined a number of threats to the validity of causal inferences that are relevant for the present research. Our study was designed to counteract four such threats that may compromise valid inferences about campaign effects. For each of the three campaigns our general approach was to compare responses given by the same persons before and after a campaign was conducted [cf. (Cook and Campbell, 1979; Kessler, 1993)].

A first potential threat to the validity of inferences about campaign effects is related to our general approach of using longitudinal data. Specifically, collecting responses by the same people at two time points may yield a testing effect; put differently, the baseline assessment may be reactive. Indeed, not only may pre-testing people affect their reception of subsequent campaign messages, but also answering questions on attitudes, intentions, norms and self-efficacy may lead to change in these variables (Cook and Campbell, 1979). To avoid this problem, in the present study people who participated both at a baseline and a post-test assessment were compared to a post-test-only control group. In the remainder of this article the two groups will be called the ‘baseline–post-test group’ and the ‘post-test-only group’, respectively. When both groups are found to have equal post-test means, baseline–post-test differences can be attributed to the campaign instead of to mere participation in the baseline assessment. Of course, this conclusion holds under the assumption that the baseline–post-test group and the post-test-only group do not differ on relevant factors, such as age, gender and sexual preference.

A second validity issue concerns the possibility of a history effect. A history effect would be present when between a baseline and a post-test assessment an event occurs that may lead to a change in the target variables. For example, a death of a celebrity due to AIDS may lead to more positive attitudes towards safer sex, particularly when this event leads to increased media attention for AIDS-related issues. In such a situation, it would be unclear whether effects are due to the campaign or to the celebrity’s death. A history effect applies to the period of time between one baseline and one post-test assessment, i.e. a history effect is only, or particularly, a potential threat to drawing valid conclusions about the effects of a campaign when only a single campaign is evaluated. Therefore, the present research examined the effects of three campaigns. When the three campaigns would show a similar pattern of results, it would seem rather unlikely that a history effect can account for the results.

A third threat concerns a cultural change, i.e. an opinion change process within society that occurs more or less independently of the campaigns that are conducted, but that might affect people’s stand concerning safer sex issues. Indeed, whereas a similar pattern of results over the three campaigns would rule out a history effect, it might still reflect a cultural change in how society as a whole perceives AIDS and STD issues. For example, attitudes might, independent of campaigns, gradually become more positive towards safer sex. To exclude a cultural change explanation, we examined whether there was a change in the target variables in the direction of a less positive stand towards safer sex in the period of time after the 1996 campaign when no campaign was conducted. If this would be the case, it would seem quite plausible that the earlier established effects reflect the effects of campaigns rather than cultural changes.

341
The last validity issue that remains is our use of a different baseline–post-test group for each of the 1994, 1995 and 1996 campaigns, and no campaign situation. Suppose that the results of each of the baseline–post-test groups would show that post-test levels differ from baseline levels in the hypothesized direction. A plausible conclusion would be that these differences reflect campaign effects, but one could also argue that differences between the various baseline–post-test groups, i.e. sample differences, caused the effects. In other words, although the use of a baseline–post-test group design for each of the campaigns allows the examination of possible campaign effects while controlling for a number of possible confounds, this procedure does not control for differences between the different samples. To optimally address our research questions, we therefore employed not only a baseline–post-test group design and a baseline–post-test group/post-test-only group design, but also a multiple assessment design. With the latter design responses given by the same people at five consecutive assessments were examined. When the different procedures would yield a similar pattern of results, sample differences could also be excluded as an alternative explanation.

The present research’s hypotheses and samples

To summarize, we tested the following hypotheses:

1. Attitudes, social norms, self-efficacy and intentions would become more favorable towards safer sex in response to each of the 1994, 1995 and 1996 campaigns.

2. In the period of time after the 1996 campaign when no campaign was conducted (but in a similar time interval as between the 1994, 1995 and 1996 campaigns), attitudes, social norms, self-efficacy and intentions would become less favorable towards safer sex.

An overview of the samples used to examine testing effects and to verify the above hypotheses is displayed in Table I. Concerning our first hypothesis, Table I shows that the first four assessments were scheduled such that between two consecutive assessments a campaign was conducted. To test our second hypothesis, the last two assessments were scheduled such that no campaign was conducted between assessments. The 1997 campaign was conducted shortly after our last assessment, but this campaign is beyond the scope of the present research.

Methods

Participants and data collection procedures

All samples consisted of participants who were selected from the households administered by the CenterData Foundation (formerly known as the Telepanel Foundation). In each of these households there is a computer that is connected with the research institute. Computers are either owned by the respondents or given on loan by the panel research institute. Each weekend participants answer a number of questions on this computer on a variety of topics. The panel is representative of the Dutch population in terms of social-demographic characteristics, which is ensured by matching the panel with data from the Dutch Central Statistical Office (CSO).

Panel members are recruited in three steps. First, a telephone survey among a random sample from the Dutch phone directories is employed to register people who are interested in participating in the panel. These people are then mailed a form that serves to assess social-demographic characteristics. Third, a random selection from the people who returned the forms is matched with CSO data, allotted a personal identification number and included in the panel. Researchers can use the identification numbers only for analytic purposes, such as linking longitudinal data to the correct panel member. Agreement to participate in the panel is not correlated with personality factors such as loneliness or need for cognition. It has been somewhat harder to recruit relatively less educated people, but in the final panel the distribution of educational levels is representative of the
Table I. Overview of designs and samples

<table>
<thead>
<tr>
<th>Sample</th>
<th>Assessment time points and campaigns</th>
<th>Sample size (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>assessed at baseline</td>
<td>post-test</td>
</tr>
<tr>
<td>2</td>
<td>assessed at</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>assessed at</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>assessed at</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>assessed at</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>assessed at</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>assessed at</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>assessed at</td>
<td></td>
</tr>
</tbody>
</table>

T1 = May 1994, T2 = June 1995, T3 = November 1995, T4 = October 1996, T5 = June 1997. Within samples, post-tests are compared to baseline assessments. Between samples, post-tests assessed at the same time point are compared to each other (except for sample 9).

Public safer sex campaigns

Dutch population. Upon becoming a member, panel members consent to participate in research on unknown topics, thus preventing systematic dropout. Participants are given the opportunity to comment on each questionnaire or not to fill out a particular questionnaire. In effect, the small non-response which potentially occurs reflects panel members who are not at home during the weekend in which a survey is administered.

For the present research, only panel members aged 15 through 45 years were sent an electronic questionnaire, since the issue of STD was assumed to be most relevant for this age group. The electronic questionnaires were programmed such that each question had to be answered before the participant could proceed, thus minimizing missing data. Whereas ‘not applicable’ was optional to answer closed-ended questions, there were no missing data concerning the target variables of the present research.

The number of waves at which each panel member participated in our research ranged from one to all five waves, which made some participants eligible for more than one sample. We therefore used personal identification numbers to ascertain that panel members were included in one sample only. Because of this procedure the samples used in the present research differ somewhat in size.

Panel members could not choose the sample (post-test-only, baseline–post-test group or multiple assessments group) in which they would be included. Thus, differences between samples could not be due to selection effects. Table II displays features of all samples. Note that there are minor differences between the samples, most notably between sample 8 and the remaining samples, concerning participants’ age, gender, educational level, relationship type and sexual preference. Since controlling for these variables did not affect our results significantly, they will not be mentioned further. The results that are presented in the remainder of this article pertain to the original, unadjusted values of our target variables.

Campaign descriptions

The 1994, 1995 and 1996 campaigns shared the general objective of putting safer sex on the public agenda. In addition, each campaign had specific goals. The 1994 campaign specifically aimed to positively affect people’s (perceived) skills to ensure that condoms are used in potential risky sexual encounters (i.e. self-efficacy). For example, in television spots couples in an intimate situation explicitly showed a condom followed by the central slogan ‘I have Safe Sex or No Sex’. All campaigns were based on the principle that whereas safer sex
Table II. Descriptive information

<table>
<thead>
<tr>
<th>Variables</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Mean age (years)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>33.6</td>
</tr>
<tr>
<td>Gender (%)</td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>55.1</td>
</tr>
<tr>
<td>female</td>
<td>44.9</td>
</tr>
<tr>
<td>Educational level (%)</td>
<td></td>
</tr>
<tr>
<td>vocational training or lower</td>
<td>41.2</td>
</tr>
<tr>
<td>high school/community college</td>
<td>30.3</td>
</tr>
<tr>
<td>college/graduate education</td>
<td>28.5</td>
</tr>
<tr>
<td>Sexual relationship (%)</td>
<td></td>
</tr>
<tr>
<td>one steady partner</td>
<td>73.6</td>
</tr>
<tr>
<td>none</td>
<td>16.1</td>
</tr>
<tr>
<td>casual only</td>
<td>10.3</td>
</tr>
<tr>
<td>Sexual preference (%)</td>
<td></td>
</tr>
<tr>
<td>heterosexual</td>
<td>96.1</td>
</tr>
<tr>
<td>homosexual/bisexual</td>
<td>3.9</td>
</tr>
<tr>
<td>N</td>
<td>441</td>
</tr>
</tbody>
</table>

For samples in which the questionnaire was administered more than once, the displayed variable means and frequencies pertain to data collected at a first assessment.

should be a prevailing norm, a moralistic approach should be avoided (Brandt, 1988). Further, STD prevention should be reconcilable with a positive approach of sexuality and campaign messages should be understandable for all [see (De Vroome et al., 1994)]. Television, radio, cinemas, posters in public places, newspapers and other periodicals, and free postcards and brochures were used to distribute campaign messages.

Important objectives of the 1995 campaign were to positively affect social norms towards safer sex and to enhance people’s self-efficacy. The campaign underscored the importance to differentiate between practical and communicative skills concerning safer sex. For example, brochures and leaflets presented explicit illustrations about how condoms should be used, whereas posters showed one-liners in text balloons suggesting what to say to ensure condom use. Examples of one-liners are ‘Would you like me to wrap it up for you...?’ and ‘If you’ll put something on, I’ll put something off...’. In addition to these one-liners, all campaign material featured the central slogan ‘I have Safe Sex or No Sex’.

Similar to the 1995 campaign, the 1996 campaign focused on social norms and practical and communicative skills (self-efficacy) concerning condom use. The campaigns were also alike in showing couples in a sexual situation, providing printed suggestions about, for example, how to bring up condom use and the central slogan featuring in all campaign material. The 1995 and 1996 campaigns differed in the colors and one-liners that were used. In the 1995 campaign bright colors were used, whereas the 1996 campaign featured more delicate colors. Examples of one-liners that were used in 1996 are ‘Your condom or mine?’, and ‘Panties off, jacket on’.

Measures of target variables

The measures were embedded in a longer questionnaire. For all constructs scales were computed by averaging scores across items. Between samples and constructs, Cronbach’s α values ranged from 0.60 to 0.92. Except for attitudes, the same measures were used in all five assessments.
To assess attitudes towards safer sex, participants were asked to indicate how they felt about using condoms with a new sexual partner (1 = unwise, 5 = very sensible) and discussing condom use with a new sexual partner (1 = unwise, 5 = very sensible). From the second assessment on, two additional items were used asking how the participants felt about discussing and using condoms with a new sexual partner (1 = not positive at all, 5 = very positive).

Social norms towards safer sex were measured with five Likert-type items. Specifically, with two items it was asked whether people who are important to the participants would approve of the latter discussing condoms (1 = not at all, 5 = very much) and using condoms (1 = not at all, 5 = very much). With two other items, participants were asked whether their friends and acquaintances would discuss condoms (1 = not at all, 5 = very much) and use condoms with a new partner (1 = not at all, 5 = absolutely). Finally, participants were asked to indicate how they thought a new sexual partner would feel about using a condom (1 = very negatively, 5 = very positively).

Self-efficacy
Four Likert-type items were used to assess self-efficacy with respect to safer sex. The first two items assessed participants’ perceptions of the difficulty of discussing and using condoms (1 = very difficult, 5 = very easy), whereas two other items assessed to what extent the participant thought (s)he could ensure that condoms were discussed and used (1 = not at all, 5 = absolutely).

Intentions
We measured safer sex intentions with two Likert-type items. The items assessed whether participants thought they would use condoms with future sexual partners and whether they thought they would discuss condom use with a future sexual partner (1 = not at all, 5 = absolutely).

The results are presented in three sections. The first section examines whether post-test scores on the target variables were independent of participation in the baseline assessment, i.e. whether a testing effect can be ruled out as an explanation of the results of our main analyses. The second section investigates changes over time by examining baseline–post-test differences in the target variables. Note that this section not only examines campaign effects, but also deals with history effects and cultural change effects as threats to the validity of causal inferences about campaign effects. In the third section changes over time are investigated by analyzing developments within a sample of persons who participated in all five assessments. With this last section we examine campaign effects and test whether the results are independent of sample differences as used in the second results section.

Testing effects
For the purpose of examining whether changes over time reflect testing effects rather than campaign effects, for each of the four sets we tested whether baseline–post-test groups and post-test-only groups had equal post-test scores on the four target variables. Out of the resulting 16 analyses of variance that we conducted, only in two cases did baseline–post-test groups differ from post-test-only groups. At the second of our five assessments, attitudes in the baseline–post-test group were somewhat more positive (M = 4.57) than in the post-test-only group (M = 4.66; F(1,1,500) = 5.13, P < 0.05). At the fifth assessment, intentions in the baseline–post-test group were less strong (M = 4.24) than in the post-test-only group (M = 4.41; F(1,5,000) = 4.69, P < 0.05). All other differences between baseline–post-test and post-test-only groups were non-significant. On the basis of these results, we conclude that there is little evidence that our assessments were affected by testing effects.

Changes over time: baseline–post-test differences
The four baseline–post-test groups were submitted to four repeated measures MANOVAs with time
Table III. Changes over time: means and SD of target variables at baseline (BL) and post-test (PT)

<table>
<thead>
<tr>
<th>Variable</th>
<th>1994 campaign</th>
<th>1995 campaign</th>
<th>1996 campaign</th>
<th>No campaign</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BL PT P&lt;</td>
<td>BL PT P&lt;</td>
<td>BL PT P&lt;</td>
<td>BL PT P&lt;</td>
</tr>
<tr>
<td>Attitudes</td>
<td>4.59 4.57 –</td>
<td>4.59 4.62 –</td>
<td>4.55 4.66 0.09</td>
<td>4.59 4.42 0.001</td>
</tr>
<tr>
<td>Social norms</td>
<td>3.97 4.08 0.001</td>
<td>4.08 4.18 0.001</td>
<td>4.08 4.19 0.09</td>
<td>4.20 4.12 0.05</td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>3.92 4.01 0.01</td>
<td>4.02 4.05 –</td>
<td>3.99 4.02 –</td>
<td>4.03 3.95 0.09</td>
</tr>
<tr>
<td>Intentions</td>
<td>3.99 4.08 0.05</td>
<td>4.20 4.21 –</td>
<td>4.41 4.36 –</td>
<td>4.19 4.24 –</td>
</tr>
</tbody>
</table>

P< denotes significance level of post-test—baseline differences; only P < 10% displayed.

of assessment (Time: baseline, post-test) as within-subjects factor and the four target variables as measures. Means and SDs of the target variables in the four groups are presented in Table III. (The small SDs and high mean values presented in Table III indicate small variances. An inspection of the data showed that in almost all cases the range of scores included the anchors of the measurement scales, suggesting non-normal distributions. Because non-normality does not seriously affect type I error and power in MANOVA procedures, we felt that our employment of MANOVAs was appropriate (Stevens, 1992).) When there would be positive changes with respect to safer sex in the target variables in response to all four campaigns, it would seem unlikely that these changes were due to history effects, and when the target variables would change in a less positive direction when no campaign was conducted, a cultural change explanation of the observed effects would seem not very likely.

The 1994 campaign
Time showed a significant effect at the multivariate level, F(4,437) = 5.84, P < 0.001. Hence, we investigated the univariate results. The baseline—post-test differences for social norms, self-efficacy and intentions were significant, F(1,440) = 14.88, P < 0.001, F(1,440) = 6.63, P < 0.01 and F(1,440) = 3.95, P < 0.05, respectively. As expected, all these differences reflected higher post-test scores than baseline scores and were thus more favorable with respect to safer sex. Changes in attitudes were not significant.

The 1995 campaign
The MANOVA yielded a multivariate effect of Time, F(4,712) = 6.11, P < 0.001. The univariate results showed a significant change in the direction of social norms more positive towards safer sex, F(1,715) = 20.28, P < 0.001. For attitudes, self-efficacy and intentions, changes over time were not significant.

The 1996 campaign
The results show that the multivariate effect of Time did not reach significance, F(4,97) = 1.53, NS. The non-significance of the multivariate effect may be explained by the small sample size. The univariate results suggest that compared to baseline assessments, levels of attitudes and social norms were somewhat more positive after the 1996 campaign, F(1,100) = 3.08, P < 0.09 and F(1,100) = 2.94, P < 0.09, respectively. However, it should be noted that the results are not particularly strong.

No campaign
In contrast with the previous three analyses, it was expected that between T4 and T5, the levels of the
target variables would become less positive. The MANOVA that was employed to test this expectation yielded a multivariate effect of Time, \( F(4,284) = 6.55, P < 0.001 \). The univariate results showed that the differences between baseline and post-test were significant for attitudes, \( F(1,287) = 15.49, P < 0.001 \), and social norms, \( F(1,287) = 4.73, P < 0.05 \). The means show that at T5, reported levels were indeed less positive than at T4. For self-efficacy and intentions, changes over time were not significant.

In general, these results show that the target variables stayed the same or became more positive towards safer sex after the 1994, 1995 and 1996 campaign. In addition, all the target variables, with the exception of intentions, became less positive towards safer sex during the period in which no campaign was conducted. This pattern of results seems to go against an explanation of the results in term of a history effect or cultural change.

**Changes over time: differences between multiple assessments**

To exclude an explanation of the findings in terms of sample differences, the multiple assessment group, i.e. the group that participated in all five assessments, was submitted to a doubly multivariate repeated measures analysis with time of assessment (Time: T1, T2, T3, T4, T5) as within-subjects factor and the four target variables as measures. Recall that at T1, attitudes were measured with two items, whereas at T2 through T5 attitudes were measured with four items. To enable meaningful comparison of contrasts with regard to attitudes, we chose to use the same two attitudes measures for all five levels of Time.

We expected the target variables to become more favorable towards safer sex over the first four assessments, and to become less favorable between the fourth and fifth assessment. Therefore, reversed Helmert contrasts were tested for T1 through T4, and a repeated contrast was tested for the difference between T4 and T5. A reversed Helmert contrast compares the mean of each level to the mean of all previous levels. A repeated contrast compares the mean of each level to the mean of the subsequent level. The overall MANOVA yielded a significant multivariate effect of Time, \( F(16,129) = 3.09, P < 0.001 \). All univariate tests of Time were also significant; for attitudes, \( F(4,576) = 6.23, P < 0.001 \), for social norms, \( F(4,576) = 5.03, P < 0.001 \), for self-efficacy, \( F(4,576) = 2.47, P < 0.05 \), and for intentions, \( F(4,576) = 5.96, P < 0.001 \).

To get a clear picture of changes over time, mean scores at each time point are graphically displayed in Figure 1. As observed earlier, all means are highly positive. Consequently, the differences between the means over time are small. Because Figure 1 serves the purpose to highlight the trends over time and not to mark out the means on the entire measurement scale, the y-axis in Figure 1 corresponds with only the upper part of the five-point scale used to measure the target variables. The trends that can be observed in Figure 1 show that in line with our expectations, all target variables became more favorable towards safer sex until T4, i.e. during the period that campaigns were conducted, whereas the target variables became less positive after T4; in effect, when no campaign was conducted. The results of the contrast analyses were examined to test this tentative conclusion.

For all target variables, four contrasts were analyzed. Specifically, the first contrast tested \( M_{T2} - M_{T1} \), the second contrast tested \( M_{T3} - (M_{T1} + M_{T2})/2 \), the third contrast tested \( M_{T4} - (M_{T1} + M_{T2} + M_{T3})/3 \) and the fourth contrast tested \( M_{T5} - M_{T4} \). The results of the contrast tests will be discussed for each target variable separately.

**Attitudes**

The first two contrasts were not significant. The test of the third contrast showed significantly more positive attitudes at T4 as compared to the mean of T1 through T3, \( F(1,144) = 6.48, P < 0.05 \). The decrease in positive attitudes between T4 and T5 was substantial and significant, \( F(1,144) = 21.06, P < 0.001 \).

**Social norms**

The reported levels of social norms did not differ between T1 and T2. Social norms at T3 were significantly more positive, however, compared to
Fig. 1. Changes over time in attitudes, social norms, self-efficacy and intentions.

The present study evaluated the 1994, 1995 and 1996 Dutch safer sex campaigns in their effectiveness of positively changing attitudes, social norms, self-efficacy and intentions.
Public safer sex campaigns

norms, self-efficacy and intentions. The design of the present study was unique in the sense that four different alternative explanations for the effects that were found could be excluded to a considerable degree. First, support was found for the adequacy of our approach of examining baseline–post-test differences by showing that a group that was administered only a post-test did not differ from a group that was administered a pre-test assessment as well. Thus, testing effects do not seem to be responsible for the effects that we found. Second, our study was not limited to the effects of a single campaign—we studied the effects of three subsequent campaigns. Given our finding that the campaign effects were highly similar across the three campaigns, a history effect seems rather unlikely as an explanation of the results. Third, we were able to show that the trend of more positive responses to safer sex was not the result of a cultural change within the society as a whole. Indeed, in the time period in which no campaign was conducted, most responses to safer sex became less positive. Finally, our study included a longitudinal sample consisting of persons who participated in all five assessments. Because this sample yielded results that were highly similar to the results for the baseline–post-test groups, we can conclude that the results were not due to possible differences between the baseline–post-test groups. By counter-acting four different threats to the validity of causal inferences, we feel reasonably confident in interpreting the present results as effects of the mass media campaigns.

Let us now look briefly at the results. The high levels of the target variables are noteworthy. Already before the beginning of campaigns evaluated here, attitudes, social norms, self-efficacy and behavioral intentions with respect to safer sex were very positive. As these variables are important predictors of safer sexual behavior [e.g. (Reinecke et al., 1996)], it is encouraging that very high levels of these variables were found. On the other hand, these high levels seem to leave little room for a further increase. Given the high baseline levels, if anything could be expected, it would be a decrease. Yet, both the baseline–post-test and the longitudinal or multiple assessment analyses showed that from T1 to T4, i.e. in the period in which the campaigns were conducted, a positive trend in the target variables could be observed. More specifically, attitudes were positively affected by the 1995 and 1996 safer sex campaigns, and the 1994 and 1995 campaigns successfully enhanced the perceived social norms towards safer sex. These results are encouraging in light of recent empirical studies that show that social norms are highly predictive of safer sex intentions and behavior among several populations [e.g. (Morrison et al., 1995; Buunk et al., 1998)]. Our findings further suggest that self-efficacy was enhanced by the 1994 campaign, and that intentions were enhanced by the 1994 and 1995 campaigns. To summarize, all target variables were generally more positive with respect to safer sex after a campaign than before.

Our finding that the target variables became less positive with respect to safer sex in the period in which no campaign was conducted supports the reasoning that the positive changes in the target variables between T1 and T4 reflect indeed effects of the campaigns. Between T4 and T5, i.e. in the period in which no campaign was conducted, the positive trend that was observed between T1 and T4 was reversed. Thus, we found that in the absence of a campaign, attitudes and social norms became less positive with regard to safer sex. Similar changes were observed for self-efficacy and intentions, but the magnitudes of these changes were smaller compared to the changes in attitudes and social norms. The overall picture that emerges from the findings is that mass media campaigns are effective in stabilizing or even enhancing the determinants of safer sex behavior and that the continuation of such campaigns is necessary to maintain high levels of these determinants. This conclusion is in agreement with evaluation results of German safer sex campaigns, that showed that the discontinuation of campaign efforts due to budget cuts led to a decrease in safer sex communication and possibly to a decrease in safer sex behavior (Müller et al., 1998). The similarity of the German results and the current study’s results
M. C. Yzer et al.

contribute to our understanding of the potential impact of safer sex campaigns.

When researchers evaluate public campaigns, it is important that they examine the same variables as the campaigns are aiming to change. In the present study we therefore focused on those determinants of safer sex that were addressed by the Dutch safer sex campaigns. A disadvantage of this approach is that the present study cannot answer the question whether the observed positive changes in determinants of safer sex led to a change in safer sexual behavior. Despite the apparent importance of information about safer sexual behavior, we feel that inclusion of behavior in the present research design would not have been appropriate, because behavior change is unlikely to be induced by public mass media campaigns. Rather, a public campaign serves to create a positive climate towards safer sex, thereby facilitating small-scale tailored interventions to establish behavior change. It would therefore be more appropriate to examine behavior when evaluating small-scale interventions than when evaluating public safer sex campaigns.

The results of the present research suggest a general trend in which the target variables became more favorable towards safer sex in response to the three campaigns. Recall, however, that the campaigns particularly aimed to enhance self-efficacy (all three campaigns) and social norms (the 1995 and 1996 campaigns). For a thorough understanding of the potential of campaigns, it is clearly important to know whether these specific objectives were realized. Interestingly, additional analyses of effect sizes and the magnitudes of the contrasts, as well as analyses in which we pitted changes in self-efficacy and social norms against overall changes revealed that the effects on self-efficacy and social norms did not exceed the general campaign effects. The present data therefore offer no evidence that changes in self-efficacy and social norms were greater than the general trend that was observed. A possible interpretation of these results is that the campaign messages addressed not only self-efficacy and social norms, but unintentionally also addressed people’s attitudes and intentions. It is also conceivable that changes in social norms and self-efficacy yielded similar changes in attitudes and intentions because these variables were strongly associated with each other (zero-order correlations in the present data sets were in the range of \(0.48 < r < 0.72\)). The best explanation for our observation that no changes in any single variable exceeded the general trend may well lie in the high pre-campaign levels of the variables. Indeed, the little room there was for positive changes critically limited the feasibility of detecting differences in the magnitude of changes.

In this paper we have focused on effects that can be attributed to public campaigns aimed at fostering safer sexual behavior. Our central assumption was that changes in levels of target variables are indicative of campaign effects, under the condition that alternative explanations can be controlled for as much as possible. Whereas we found considerable evidence that the campaigns as a whole yielded effects upon the target variables, it remains essentially rather difficult to establish which specific campaign elements led to the results, because each campaign consisted of a host of different elements that were in constant interplay with each other. Nonetheless, it can be concluded that the combination of the principles on which the campaigns were founded, the message characteristics and the various media used to diffuse the messages apparently constituted an effective communication mix. An overview of campaign principles, message characteristics and media was presented in the Methods section of this paper.

It should be noted that although our research design tried to counteract a number of potential threats to the validity of causal inferences about campaign effects, we do not imply that our results can only be interpreted as campaign effects. Indeed, campaigns are conducted in a dynamic setting in which a multitude of factors interacts with campaign effects. Given our research design and methods of analysis, however, our results strongly suggest that the changes that we observed are at least partially due to effects of the campaigns. To enable still more unequivocal conclusions about campaign effects, future evaluation research could...
extend the design used in the present study with a measure of exposure to campaigns [e.g. (Middlestadt et al., 1995)]. This would enable a quasi-experimental design that includes an experimental group, consisting of those who were exposed to campaigns, and a control group, consisting of those who were not or to a lesser degree exposed to a campaign. Although the use of exposure measures in evaluation research is a promising concept, it may prove to be very difficult to construct a measure that is sufficiently sensitive in assessing exposure to a campaign. Indeed, there is much discussion about how to assess exposure to mass media communications (Roberts and Maccoby, 1985). The present data set included a self-report measure of exposure to the campaigns. Additional analyses showed that compared to levels of relatively weakly exposed individuals, the target variables were more positive for those who reported relatively strong exposure. However, the target variables did not become more positive for relatively strongly exposed individuals than for relatively weakly exposed individuals. We could therefore not demonstrate that the magnitude of changes in target variables differed as a function of exposure to the campaigns. This could well be due to the fact that our measure was not sufficiently sensitive in assessing exposure independent of other factors. In general, exposure measures are likely to be associated with a number of confounding factors, such as personality factors, experience with the behavior, past opinions, etc. (Roberts and Maccoby, 1985). Demonstration of direct effects of exposure on target variables is therefore rather difficult.

The findings reported in the present article lead us to elaborate upon the implications of our findings for future safer sex campaigns. The changes over time that we found suggest that the campaigns were effective in changing the determinants of safer sex. On the other hand, the levels of these determinants were already very high before any campaign was conducted. This may lead one to conclude that there is little use in conducting an expensive mass media campaign when only small changes are to be expected. However, we also found evidence that attitudes and social norms became less positive when the campaigns were discontinued. One of the assumptions behind the campaigns in the Netherlands was that a campaign should keep AIDS and STD in the public awareness (Kolker, 1997). The stabilization of the high levels of determinants of safer sex over time as found in the present study suggests that the campaigns were effective in reaching this goal. Further, we established that an absence of public campaigns negatively affected people’s attitudes, social norms, self-efficacy and intentions. Therefore, we feel that campaigns still have an important role to play. Campaign goals, however, need to be set realistically and may be formulated in terms of stabilization rather than in terms of dramatic change.

Acknowledgements

This article was supported by ‘Zorgonderzoek Nederland’ (grant 95.016). We would like to acknowledge the cooperation of the Dutch Foundation for STD Control.

References


M. C. Yzer et al.


Received on April 16, 1999; accepted on September 17, 1999