Published in final edited form as:

AIDS Educ Prev. 2009 April ; 21(2): 128–140. doi:10.1521/aeap.2009.21.2.128.

REPLICATING AN INTERVENTION: THE TENSION BETWEEN FIDELITY AND ADAPTATION

Diane M. Morrison,

School of Social Work, University of Washington, Seattle

Marilyn J. Hoppe,

School of Social Work, University of Washington, Seattle

Mary Rogers Gillmore,

School of Social Work, Arizona State University, Phoenix

Carisa Kluver,

School of Social Work, University of Washington, Seattle

Darrel Higa, and

School of Social Work, University of Washington, Seattle

Elizabeth A. Wells

School of Social Work, University of Washington, Seattle

Abstract

Increased awareness of the importance of tailoring interventions to participants' cultures has focused attention on the limited generalizability of a single test of an intervention to determine efficacy. Adaptation is often necessary to replicate interventions across cultures. This produces a tension between fidelity to the original intervention and adaptations necessary to make the intervention relevant to the culture and circumstances of participants. This article discusses issues that arise during the course of replication, with illustrations from a replication to test the efficacy of an HIV prevention intervention for youth, using a randomized controlled design. Analysis of the issues raised leads us to suggest that a "science of replication" needs to be developed.

When is an implementation of an established intervention a replication as opposed to a new intervention? In the strictest sense, any subsequent implementation of an intervention is necessarily, at least in part, a different intervention, because every situation is unique. Even when investigators repeat their own interventions in the same sites, they replicate them in an environment at least somewhat changed with the passage of time and with different participants who bring their unique characteristics to the process. Often, substantive changes are made in interventions to adapt them to a new site, to keep them up-to-date, or to tailor them to a new population. Under what conditions can we expect an intervention of demonstrated efficacy to continue to be efficacious; how far, or in what dimensions, can we change an intervention and still be confident that it is the same intervention, and will have the same effects?

Most of the literature on replication of interventions focuses on problems of translating interventions that were effective in efficacy trials to the "real world" of practice (e.g., Elliot & Mihalic, 2004; Glasgow, Lichtenstein, & Marcus, 2003). The failures of interventions that were promising in initial efficacy trials but less effective, or ineffective, in applied settings

(effectiveness trials) has prompted researchers to look at differences in control, standardization, and resources that are inherently different in the two settings (Clarke, 1995; Glasgow et al., 2003; Kelly et al., 2000; King, Hawe, & Wise, 1998; Kraft, Mezoff, Sogolow, Neumann, & Thomas, 2000; Rotheram-Borus, Rebchook, Kelly, Adams, & Neumann, 2000; Simpson, 2002; Visser & Schoeman, 2004). Relatively little, however, has been written about the issues of replication between experimental studies testing efficacy. Examining the planned and unplanned changes that arise in a replication of an experiment to test an intervention, rather than in the transfer to an applied setting, allows us to distinguish between changes that are inherent in any replication and those that are a function of a change in environment from “laboratory” to “field.” Replicating an experimental study to test efficacy is less common; typically, once an intervention has been demonstrated to be efficacious, it is assumed that it would be similarly successful were it to be tested again, under similar conditions. However, increased awareness of the importance of adapting or tailoring interventions to the culture of the participants has focused attention on the limited generalizability of a single experimental intervention to test efficacy, especially one designed and tested with participants who are homogeneous with regard to race, ethnicity, socio-economic status, or other demographic characteristics (Castro, Barrera, Martinez, 2004; Rodrigue, Tereyak, & Lescano, 1997).

Lessons learned about the importance of fitting interventions to the cultural context of the participants have raised new conundrums. Whereas the “gold standard” for a replication had been strict fidelity to the content and procedures of the original project, we now understand that slavish fidelity may result in an intervention that is faithful to the form, but not the essence, of the original. There is a tension between fidelity and adaptation that cannot be resolved easily or simply. We describe here the issues that researchers confront, with illustrations from our experimental replication study that tested efficacy in a different population, as a first step toward delineating the changes, both planned and unplanned, that arise in the process of replication. We do not focus here on the findings of the study, which have been reported elsewhere (Morrison et al., 2007). Overall, the findings were equivocal, in part owing to lower rates of sexual activity among the youth in the replication sample; there were no significant differences in cognitions or behavior but a trend in condom use similar to that reported in the original study.

THE STUDIES: FOCUS ON KIDS AND TEENS TAKE CHARGE

Both the original study (Focus on Kids) and our replication (Teens Take Charge) were conducted as randomized field experiments, funded by grants from the National Institutes of Health. In this way, they were more similar in situational control, training of staff, and resources than the laboratory-to-field, efficacy-to-effectiveness transitions that have been more widely discussed. The original investigators had published widely on the study, giving us good information on how the project was implemented, and the complete curriculum was published by ETR Associates (1998). We were very fortunate to have consultation and training from the principal investigator and project director of the initial study. Thus, the conditions for a careful replication were close to ideal.

The original intervention, Focus on Kids (FoK), is an AIDS prevention intervention developed specifically for inner-city African American youth by Bonita Stanton and her colleagues at the Center for Minority Health Research at the University of Maryland (Galbraith et al., 1996; Stanton et al., 1995; Stanton et al., 1996). It has documented efficacy and is culturally grounded in ways that experience suggests are essential to successful implementation. Its focus on community involvement and sustainability help ensure maximum participation and inclusion of the most knowledgeable community partners, who are often wary of entering into relationships with academic researchers. The intervention is grounded in theory, specifically protection motivation theory (Rogers, 1983).

FoK was developed in, and delivered in, housing project recreation centers in Baltimore, MD. The intervention was delivered to small groups of three to nine same-sex friends, ages 9 to 15, who met in eight weekly sessions. All of the sessions except Session 6 were about 1.5 hours in length. Session 6 was a daylong retreat at a rural campsite, combining youth and facilitators from all of the groups. The sessions focused on one or more of the protection motivation theory constructs. The initial session's primary goal is to build trust and group cohesion through a variety of games and exercises, but it also introduces a decision-making model, the SODA Model, which teaches a four-step process of decision-making (stop and state the problem, consider the options, decide and choose the best solution, act on the decision). Other sessions focus on risks and values, information and decision making skills, consequences of behavior, skill building, information about sexual health, and attitudes and skills for sexual health, respectively. The curriculum emphasized decision making regarding a variety of risky behaviors, including sexual behavior, substance use, and violence, and used a variety of formats to deliver content (e.g., videos, games, discussions, role-playing, etc.).

The intervention's efficacy was tested in a randomized experiment in which the *FoK* intervention was compared with a control group that met to watch films and videos about AIDS, AIDS prevention, and other risky activities. A total of 383 youth participated in the study (206 who received the *FoK* intervention compared to 177 who did not), in single-sex groups. The intervention was delivered to each group by two trained adult facilitators recruited from the community, one of whom was the same gender as the group, and most of whom were African American. The tasks and activities, materials needed, and time for each activity for each session were spelled out in detail in a manual that was followed by the facilitators. The intervention's efficacy was assessed with a computerized multicomponent risk assessment questionnaire at baseline, and again at 6, 12, 18, 24, and 36 months postintervention. The questionnaires assessed risk behaviors in the past 6 months, including sexual intercourse, unprotected intercourse, substance use (alcohol, cigarettes, and illegal drugs), violence (carrying guns, knives or bats, fighting), and truancy. Overall, the effectiveness of the intervention on condom use was positive (Stanton et al., 1996).

We replicated *FoK* in a different community to determine its generalizability to other populations. The replication project was called Teens Take Charge (TTC), and took place in Seattle, WA. It involved an ethnically more diverse population of 12–15-year-old African Americans, Asian Americans, and African and Asian immigrants. Youth in the TTC project did not necessarily live in large subsidized housing projects or in close proximity to the other youth in their intervention groups. Our goal for the replication was to stay as close as possible to the original implementation of the intervention, except insofar as necessary to tailor material to the youth in our community and to the community centers that serve them.

Similar to *FoK*, the intervention consisted of eight weekly sessions of about 2 hours in length. These sessions were followed by a 2-hour booster session at 9 months postintervention to review and reinforce the intervention material. Two trained adult facilitators who had been recruited from the community led the sessions. Facilitators followed a detailed manual, based on the *FoK* curriculum, that provides specific instructions for the tasks and activities for each session. The large majority of these sessions were identical to those originally designed for *FoK*. Like Stanton and colleagues, we assessed the effects of the TTC intervention using a computer-administered questionnaire with audio assist. The questions were similar, and sometimes identical, to those used by Stanton and her colleagues. Youth were assessed at baseline (pre-intervention), immediately following the 8 weekly sessions (posttest) and at 6- and 12-month post-intervention follow-ups.

CHANGES AND ADAPTATIONS

Below, we outline the types of modifications that may arise in a replication project, with illustrations from our replication study. They are grouped into categories of scientifically motivated changes, geographically mandated changes, changes in response to community composition and culture, changes in support of collaboration with community partners, changes in the role of facilitators, updates, and field adaptations.

SCIENTIFICALLY MOTIVATED CHANGES

When initially developing an intervention, there is scientific and ethical justification for a no treatment or minimal treatment control group. Scientifically, an initial trial is intended to establish that there is some efficacy to the intervention; a replication study, at least implicitly, begins a component analysis: identifying the elements that make the intervention efficacious. Ethically, when one does not know whether the intervention is effective, those in the control group are not a priori being deprived of a benefit. When replicating, this is no longer the case: The intervention is at least very likely to be beneficial, making it more imperative to offer something of comparable value to the control or comparison group participants.

In the original FoK, control youth viewed sex education films, followed by discussion. For the replication, we needed a control group comparable in intensity, length, and attractiveness to the intervention. Based on feedback from focus groups of youth and from input from our community advisory board, we decided on and developed a career exploration curriculum to serve as the control group. The control group sessions were of the same number and length, as the experimental intervention. Groups were formed in the same way, were of the same size, and were led by trained facilitators who followed a comparably detailed manual.

It would be blind consistency to duplicate components or aspects of an intervention that were found to be ineffective in the initial trial. For example, Stanton et al. (1996) found that the FoK intervention was effective primarily among the teens and had little effect for younger youth (9–11 years old). Accordingly, we narrowed the age range of the participants to 12 to 15. Similarly, Stanton et al. found that intervention effects observable at 6 months post intervention had begun to deteriorate by 12 months post intervention. She instituted a booster session at 15 months, which was effective in reinstating differences between the control and experimental groups at 18 months. Based on her experience, we planned a booster session to be delivered 9 months after the posttest, in an attempt to avoid the 12-month deterioration she observed.

It is also essential to assess fidelity in a replication so that if the intervention fails to replicate you can assess whether this was due to poor implementation. In TTC, this meant that we added measures of fidelity to our project protocol, which required youth and facilitators to complete a checklist about the content covered, and audiotaped all sessions, followed by auditing a sample of sessions.

GEOGRAPHICALLY MANDATED CHANGES

The physical aspects of a new community can offer unexpected challenges to faithful replication of an intervention study. Recruitment and retention can be affected by differences in the physical accessibility of the intervention. In a group intervention, geography may dictate the extent to which participants interact outside of the intervention setting. Housing patterns of Baltimore and Seattle differ greatly. Youth in the FoK study in Baltimore lived in the same large housing projects, went to the same schools, and “hung out” at the same after school venues. In Seattle, housing is much less dense, and large housing projects are rare.

Lack of density in Seattle also affected the design of the trial. In Baltimore, because large numbers of youth were concentrated in three public housing developments, it was possible to

recruit and conduct all of the groups simultaneously. In Seattle, recruitment proceeded much more slowly. No single agency or after-school program could provide all, or even a large fraction, of the number of youth we enrolled. We therefore needed to recruit groups in waves, over two years. This increased the possibility of cross-group contamination; however, because youth did not live in close proximity, this possibility was lowered.

School assignments are handled differently in different locations, with large effects on recruiting for a youth intervention. Seattle's school busing plan results in youth from the same neighborhood often attending different schools. Seattle participants would not be in the same kind of frequent contact with other members of their TTC groups; although they were served by the same agency or attended the same after school program, they might not live in the same neighborhoods or attend the same school. Components of the program that relied on proximity and interaction outside of TTC groups needed to be changed (e.g., group projects to be conducted outside of meeting times).

Another geographic factor that had a large impact was the relative heterogeneity of racial or ethnic groups in Seattle. In Baltimore, all of the participants were African-American inner-city youth, as were the great majority of facilitators. In Seattle, neighborhoods are generally less segregated, and there is a large immigrant population that lives in the same neighborhoods as many African Americans. With more diverse groups of youth, it was not possible to match ethnicity of facilitators for all youth. In our mixed-ethnicity groups, some youth saw their ethnicity reflected in their facilitators, and others did not. (We were encouraged by Jemmott, Jemmott, Fong, and McCafferee's (1999) findings that facilitator race did not affect their Be Proud, Be Responsible intervention's efficacy or the participants interest in it.)

Drawing from a wider geographic region also required that we provide transportation for some youth. This led, in turn, to a number of other adaptations. To provide predictability for parents about when their children would be transported home, we standardized the length of each session. Baseline and posttest surveys were administered at the project sites, because this was the only venue in which the youth were together; the number of sessions was increased to allow time for this, and content was moved among sessions accordingly.

Recruiting over an extended period, in turn, increased the importance of our project image in the community. The word *kids* was disliked by focus groups youth, who were, on average, older than FoK youth; "teens" was more appealing. The phrase *Take charge* fit their preference for action-oriented, powerful names. A graphic artist was hired to design a logo that represented and appealed to both genders and which included diverse ethnic groups. The final version of the logo was used for stationery and other media that helped increase recognition in the community.

CHANGES IN RESPONSE TO COMMUNITY COMPOSITION AND CULTURE

One of the important elements of most successful interventions, including FoK, is responsiveness to the community's culture. This presents a particular challenge for a replication: Is it more faithful to the intervention to import its content religiously or to adapt it to the new population? Most researchers now recognize the importance of making interventions culturally appropriate (cf, Castro et al., 2004; Mrazek & Haggerty, 1994; Rotherum-Borus et al., 2000; Stanton et al., 1995). Doing so is likely to evoke changes in the intervention, however.

The heterogeneity of Seattle youth presented a challenge to a key activity in the FoK curriculum, the "Family Tree." This exercise presented a hypothetical family of a generic type that was common to many of the Baltimore youth. Youth added names and personalities and generated stories about the family. The family was introduced in the first exercise and returned to in several subsequent sessions and was used to illustrate how life choices parents make affect

their children. In Seattle, not only was the example provided not a common family constellation, the diversity of youth meant that there was no single family type that would be appropriate for all groups. Our version, therefore, needed to be not a simple revising of the family constellation but rather an adaptation that would allow the youth in each group to construct the family structure (e.g., with parents born in the United States or foreign born). This allowed youth to invent a family that reflected their own community norms, ensuring equivalent investment in and ownership of the family. Allowing this much variation in the story added greatly to its complexity and, in many cases took longer than originally planned.

Different communities have different norms about behavior; what is standard or acceptable in one may be maladaptive or unacceptable in another. Prescriptions for behavior are particularly sensitive. This arose in an exercise about effective communication. In the FoK exercise, youth role-played interactions with a salesman that were labeled “aggressive,” “assertive,” and “nonassertive,” and the curriculum implied that indirect ways of communicating were generally incorrect or inferior. For some traditional Asian Americans and Asian immigrants, however, less direct styles of communication are the norm, and the curriculum’s preferred behavioral style (“assertive”) is often inappropriate. Our adaptation retained role-plays of these three interaction styles and discussed the appropriateness of assertiveness in different contexts, but it avoided assigning evaluative descriptors or suggesting that the “nonassertive” style was incorrect. A similar issue arose in a lesson on date rape. In the original version, the incident occurred at an unchaperoned evening party. Some Seattle teens from immigrant families were not allowed to go to unchaperoned parties or dances or to be out at night. A second version of the story was added, in which the incident takes place in the young woman’s home while her mother is at work. Facilitators chose the more appropriate version for each group.

Community norms may also vary with regard to what content it is important to include. The original FoK discussed only heterosexual contact. The omission of references to same-sex activity, although appropriate in Baltimore, where this was a more sensitive topic, was seen as problematic in Seattle. To ensure that the curriculum was inclusive of gay and lesbian youth, we made changes in wording in several lessons so that the messages could apply to either heterosexual or nonheterosexual youth, and we incorporated mentions of same-sex sexual activity where appropriate. Conversely, the FoK curriculum also addressed some other problem behaviors, including drug dealing. Drug dealing was far less prevalent among our youth, so references to drug dealing were dropped from the curriculum. In a lesson on date rape, the original story refers to typical “steps” through which such incidents escalate. These were not supported by current research and were dropped. More rape “myths” were added to a list of rape myths to be discussed by the group, along with additional discussion questions, including questions about same-sex rape and rape of boys.

Community factors may affect other elements of the design, as well as content of the intervention. The inclusion of youth whose families had recently immigrated complicated consent gathering. Our initial consent procedures were similar to those used in the FoK study: Youth took consent forms home to be signed by their parents. However, an early incident made it apparent that not all youth were accurately describing the project to their parents, characterizing it as an after-school job or other activity. Some parents had trouble with the written information we sent home, even when it was translated into their native language. This led to hiring consent gatherers, fluent in the parent’s preferred languages, who phoned or visited the parents after consent forms were received. They made sure that parents understood the nature of the program and allowed them to get answers to any questions. Parents were able then to withdraw their children from the program if it was not as they had understood, and some did.

CHANGES IN SUPPORT OF COLLABORATION WITH COMMUNITY PARTNERS

Recruiting and conducting this intervention, as with many community interventions, hinged on the support and help of community agencies. Both FoK and TTC were conducted with community agency partners. The needs of community partners may differ (both between communities and within communities, over time), requiring adaptations in the research design and curriculum.

The needs and constraints of agency resources place bounds on how an intervention can be implemented. An important element of the original Baltimore FoK was drawing primarily from (and paying) staff in the agencies where the intervention was delivered. In Seattle, this was not feasible in most cases. Agency staff were generally not available for training or for assignment to the intervention for extended periods, and agencies preferred that we use staff hired by the project. Instead, we hired a core of staff facilitators who were paired with hourly facilitators to conduct sessions. This had a number of repercussions. Unlike the FoK facilitators, who typically had ongoing relationships with the young people in their groups, TTC facilitators did not have continued contact with youth between sessions. The effects of such a change are hard to gauge. This may have diminished the rapport between facilitators and participants. Conversely, it may have had a beneficial side effect of making TTC youth more comfortable asking sensitive questions of facilitators.

We found that community agencies who served our initial target populations, Asian Americans and African Americans and children of southeast Asian immigrant parents, also served increasing numbers of East African immigrants. They asked that we not exclude these youth from our study. Including African immigrant teens led to the incorporation of yet another perspective into the intervention (e.g., into the family tree), and making adjustments such as accommodating religious holidays in our schedule, and not serving snacks until after dusk during Ramadan.

Some challenges to the curriculum came from the agency sites. The curriculum calls for several activities that involve youth-handling condoms. At a few sites, the agency placed limitations on condom distribution, so alternatives were devised. For the condom demonstration (in which youth put condoms on wooden models) and the “condom race” (a team competition to demonstrate correct condom use quickly), an alternative activity had youth putting into order a set of cards that describe the steps of using a condom correctly.

CHANGES IN THE ROLE OF FACILITATORS

The role of facilitators in initial and replication studies may be profoundly different. In an initial implementation of an intervention, facilitators may have a voice in the design and implementation of the curriculum as well as leeway to adapt or drop components that they do not feel are useful. They play a creative role, helping to devise ways to translate the investigators’ ideas into a workable program, and their expertise and experience are evident and valued. In a replication study, staff are actively discouraged from making changes; fidelity to the scripted program may be monitored, with deviations prompting retraining. This can lead to profound differences in facilitator feelings of ownership of, and commitment to, the intervention.

At the outset of our project, research and field staff met regularly over several months to plan adaptations of the curriculum to the Seattle community and to a greater diversity of target youth. It quickly became apparent that “adapting” and “tailoring” had different meanings to research and field staff. Researchers planned to change only those elements of the curriculum that were poorly aligned with the planned target communities. Differentiating between tailoring

and improvements entailed extended dialogues between field and research staff, and agreements were ultimately reached.

In TTC, this affected training and supervision of facilitators, which focused on developing an understanding of and commitment to fidelity. More training than we had expected was required to achieve “buy-in” to the need for fidelity and a sense of the larger goal of the project. Emphasis was on covering all the topics in each session, meeting the stated objectives of the session, keeping the group on topic, providing the same length and intensity of contact for each group, and conducting the activities as closely to the curriculum as possible. We added a process evaluation to collect input from facilitators on different aspects of implementation.

UPDATES

Interventions require ongoing small changes in response to new information or events. A group problem-solving activity, “Burning Buildings,” used the imagery of firemen rescuing citizens from the twin towers of the World Trade Center, which we felt was inappropriate after the events of 9/11. Instead, we chose “Climbing Mt. Rainier,” where the same task was accomplished with imagery of negotiating a crevasse. New movies about puberty replaced those used in FoK, and information about available methods of contraception was updated. Changes were made to keep current on recent research in HIV prevention and treatment. During the time that our project took place, the FDA approved of several new birth control methods, access to Norplant ended, and research increased doubts about spermicidal use of Nonoxynol 9.

FIELD ADAPTATIONS

Probably the biggest instigator of ongoing adaptations is time constraints caused by unexpected problems. Inevitably, problems arose in groups stemming from disruptive behavior, interpersonal conflict, ethnic animosity or discomfort, group cohesion, and attendance problems. Youth were divided into same-sex groups, which affected the levels of disruptive behavior across groups (boys were generally more disruptive than girls), and facilitators’ ability to complete activities within the available time. All groups had times when they became engaged in important conversations—on the topic, but exceeding the time allowed and so reducing time allotted to other activities. In a few groups, internal battles between youth participants resulted in serious disruption—delaying the completion of activities and sometimes requiring additional follow-up, including parent and/or school contacts. In groups of young adolescents, a completely uneventful session can be the exception rather than the norm. Ongoing adaptations to the curriculum in response to problems are potentially as important and numerous as those made thoughtfully in research meetings but often with less documentation and attention. These unscheduled adaptations are the concern of traditional assessments of program “fidelity.”

In the field, when problems arise that result in too little time to complete the intervention as scripted, facilitators may rely on “implicit theories” about what elements of the intervention are core or are most effective (Binson, Woods, Ekstrand, Freedman, & Galvez, 2002). In the absence of a formal component analysis, facilitators fall back on their intuitions about what components are essential and need to remain exactly the same, and which are modifiable and less essential to the curriculum and to the intervention as a whole. This results in uncontrolled variation in the implementation of the intervention. If the uncontrolled variation introduces random error, the study will need more power to detect effects; if it introduces systematic error, estimation of the intervention’s effectiveness is likely to be biased. Strategies to reduce this source of uncontrolled variation used in intervention studies include careful training of facilitators, a detailed written manual with all activities clearly described, careful supervision of facilitators, and opportunities for retraining if facilitators drift from the scripted curriculum.

Despite these guards against facilitator-introduced uncontrolled variation, we suspect that such variation will never be possible to completely control. This problem arises in all field interventions, not just replication studies, but may be exacerbated by lower facilitator “buy-in.” They may result in differential delivery in replication versus original research if they arise with different frequency or at different points in the intervention.

DISCUSSION

We made adaptations in our replication of the FoK intervention at several points: preimplementation, including protocol changes (e.g., narrowing the age range of participants), and curricular changes (e.g., updating the information on HIV/AIDS); at the outset of implementation at particular sites (e.g., offering different options for a condom exercise for sites that were uncomfortable handling condoms); and as the study progressed, in response to unanticipated issues (e.g., providing transportation). With each adaptation, we experienced a tension between trying to faithfully replicate the FoK intervention while at the same time making necessary adaptations for the more diverse groups of youth who participated in our TTC study. On the one hand, an exact replication allows one to determine more definitively (other things being equal) whether an intervention’s effects are generalizable. On the other hand, failing to adapt the intervention to the new setting and more diverse youth would surely lessen its effectiveness, making replication an unfair test of generalizability. Any adaptation raises the methodological problem that the interventions are no longer identical. Thus, should it occur, a failure to replicate effects could be due to changes made in adapting the intervention for a new population. Most interventions contain many components, and we know little about which components are crucial to the intervention’s effectiveness. Few empirical studies have been conducted to examine these issues, and we lack any commonly accepted conceptual framework for approaching this question.

CONDUCTING COMPONENT ANALYSES

Ultimately, controlled studies to determine which components of an intervention are essential can provide information to use in training facilitators to deliver truly critical components of the intervention faithfully. Moreover, such information can help us design more efficient interventions that are less costly in terms of time and resources. Although component analyses are easy to design in theory, however, they can be complicated to carry out in practice (Vincent et al., 2000). There are two major problems researchers face when attempting such analyses. The first is how to define a “component.” Is a component a single session in a multi-session intervention? Is it all the content of a certain type in the intervention regardless of whether it is spread out over several sessions (e.g., negotiation skill development)? Is a component all the content on a single construct from the theory guiding the intervention, regardless of whether it occurs in a single or multiple sessions (e.g., changing perceived norms about condom use)? Some initial consensus on how to define a component is needed. Content focused on the constructs of the theory guiding the intervention may be a wise starting point. The second problem is the feasibility of performing such an analysis. One might imagine it as an analysis-of-variance-type problem. With only two components (A & B), one would need to randomize study participants to three conditions (A, B, AB). With three components (A, B, & C), however, one would be randomizing participants to seven conditions (A, B, C, AB, AC, BC, ABC). With additional components, the study design becomes increasingly unwieldy. Thus, it is essential to begin a component analysis with a strong hypothesis about which elements are, indeed, core.

How can one develop these hypotheses? Kelly and colleagues (2000) have suggested that core elements can be identified through behavioral science theory, or through experience implementing an intervention. Many effective HIV prevention interventions, for example, have been based on theories drawn from the family of social cognitive theories (e.g., social learning

theory [Bandura, 1986, the theory of reasoned action [Fishbein & Ajzen, 1975], etc.). In these theories, constructs such as knowledge and skills are assumed to act on attitudes, social norms, and self-efficacy, leading to changes in intention and, ultimately, behavior. Kelly et al. also suggested that feedback from facilitators and participants can be useful in assessing core components.

ATTENDING TO PROCESS

A useful tool for research replications would be ongoing monitoring procedures to “check in” on interventions as they evolve. This would allow deviations from the protocol to be used as (admittedly imperfect) “natural experiments” in determining key components of the intervention. Post-hoc comparisons of those among whom the intervention was effective vs. those among whom the intervention was ineffective can provide hypotheses about core elements. If the intervention was less effective among those who missed a particular session, the material in that session may potentially be essential. Because missing particular sessions is not randomly assigned, the results of such post-hoc analyses are of course not definitive, but data from experimental trials can and should be mined for suggestions about where attention might be profitably focused in future studies.

USING UNEXPECTED FINDINGS

Unexpected or hard-to-interpret results may also lead to intriguing hypotheses about core elements. For example, in the Choices Project (Baker et al., 2003), a 16-week HIV/STD prevention intervention for high-risk adult women, women in the experimental group had fewer new STD infections in the year following the intervention than did women in the control group, despite the absence of significant differences in condom negotiation skills or number of risky sexual acts. The presence of experimental effects in the absence of differences in presumed core elements led the researchers to focus on potentially important variables that were *not* included in measurement (e.g., selection of less risky sexual partners or accurate condom use) as potential core elements and factors to be attended to in future studies. Similarly, “natural experiments,” in which a potentially key element changes during the course of the intervention (e.g., a particular activity coming at the end of a long session was frequently dropped) can be examined for leads about what elements are and are not essential.

CONCLUSIONS

We suspect that our experience replicating an intervention efficacy test is not unlike that of others who replicate studies or who conduct effectiveness trials. Clearly there is a need for a “science of replication”; procedures for documenting intervention content, delivery, target populations, successes, and challenges; and tools for identifying core elements. Although replication is an acknowledged component of intervention development, it is rarely discussed in the scientific literature. In the report of the Institute of Medicine presenting the “preventive intervention research cycle,” the third step includes confirmatory and replication trials of the intervention, but tellingly, in the publication of this report there is no discussion of replication techniques or challenges (Mrazek & Haggerty, 1994). The level of detail required is rarely found in published studies, no doubt in part because of publisher constraints on article lengths. Internet, rather than print, journals, or Internet links to more detailed information about published studies, are potential solutions to this problem. A second reason that this level of detail is often not found is that it adds considerable expense to a project to reliably document this level of detail when one is launching a multicomponent intervention. The solution to this problem is more complex but would be greater aided by demands for such information, along with the additional resources necessary, from the funding agencies.

Another disincentive may be the reluctance of researchers to divulge full information about intervention delivery for fear that the study in which the researcher has invested so much time and energy will be regarded as methodologically flawed, even if the original study design was methodologically sound. The solution to this problem is the most complex because it involves changes in norms in the scientific community. This is not to suggest that seriously flawed studies should be published, but rather that no in vivo intervention project will be without surprises and errors and that being frank about our unfortunate experiences can add to our cumulative knowledge, greatly strengthening future intervention studies. Developing a “science of replication” can begin with basic research on the natural history of interventions in the field: how they evolve over time, what sorts of mutations are ultimately beneficial, and what sorts are harmful. This will lay the groundwork for studying what is effective in enhancing fidelity to the intervention as scripted, and how to planfully implement cultural adaptations that retain the “active ingredients” of the intervention.

Acknowledgments

Grant Number HD38420 from the National Institute of Child Health and Human Development provided support for this research.

References

- Baker SA, Beadnell B, Stoner S, Morrison DM, Gordon J, Collier C, et al. Skills training versus health education to prevent STD/HIV in heterosexual women: A randomized controlled trial utilizing biological outcomes. *AIDS Education and Prevention* 2003;15:1–14. [PubMed: 12627740]
- Bandura, A. *Social foundation of thought and action: A social cognitive theory*. Englewood Cliff, NJ: Prentice-Hall; 1986.
- Binson, D.; Woods, B.; Ekstrand, M.; Freedman, B.; Galvez, S. *Prevention Newsletter*. Vol. 11. San Francisco: Center for AIDS Prevention Studies; 2002. What providers think about HIV prevention: The implicit theory project.
- Castro FG, Barrera M Jr, Martinez CR Jr. The cultural adaptation of prevention interventions: Resolving tensions between fidelity and fit. *Prevention Science* 2004;5:41–45. [PubMed: 15058911]
- Clarke GN. Improving the transition from basic efficacy research to effectiveness studies: Methodological issues and procedures. *Journal of Consulting and Clinical Psychology* 1995;63:718–725. [PubMed: 7593864]
- Elliot DS, Mihalic S. Issues in disseminating and replicating effective prevention programs. *Prevention Science* 2004;5:47–53. [PubMed: 15058912]
- ETR Associates. *Focus on kids: Adolescent HIV risk prevention*. Santa Cruz, CA: ETR Associates; 1998.
- Fishbein, M.; Ajzen, I. *Beliefs, attitudes, intentions and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley; 1975.
- Galbraith J, Ricardo I, Stanton B, Black M, Feigelman S, Kaljee L. Challenges and rewards of involving community in research: An overview of the “Focus on Kids” HIV risk reduction program. *Health Education Quarterly* 1996;26:383–394. [PubMed: 8841821]
- Glasgow RE, Lichtenstein E, Marcus AC. Why don’t we see more translation of health promotion research to practice: Rethinking the efficacy-to-effectiveness transition. *American Journal of Public Health* 2003;93:1261–1267. [PubMed: 12893608]
- Jemmott JB III, Jemmott LS, Fong GT, McCaffree K. Reducing HIV risk-associated sexual behavior among African-American adolescents: Testing the generality of intervention effects. *American Journal of Community Psychology* 1999;27:161–187. [PubMed: 10425698]
- Kelly JA, Heckman TG, Stevenson LY, Williams PN, Ertl T, Hays RB, et al. Transfer of research-based HIV prevention interventions to community service providers: Fidelity and adaptation. *AIDS Education and Prevention* 2000;12(Suppl A):87–98. [PubMed: 11063072]
- King L, Hawe P, Wise M. Making dissemination a two-way process. *Health Promotion International* 1998;13:237–244.

- Kraft JM, Mezoff JS, Sogolow ED, Neumann MS, Thomas PA. A technology transfer model for effective HIV/ AIDS interventions: Science and practice. *AIDS Education and Prevention* 2000;12(Suppl A): 7–20. [PubMed: 11063066]
- Morrison DM, Hoppe MJ, Wells EA, Beadnell BA, Wilsdon A, Higa D, et al. Replicating a teen HIV/ STD preventive intervention in a multi-cultural city. *AIDS Education and Prevention* 2007;19:258–273. [PubMed: 17563279]
- Mrazek, PJ.; Haggerty, RJ. Reducing risks for mental disorders: Frontiers of preventive intervention research. Washington, DC: National Academy; 1994.
- Rodrigue, JR.; Tercyak, KP., Jr; Lescano, CM. Health promotion in minority adolescents: Emphasis on sexually transmitted diseases and the human immunodeficiency virus. In: Wilson, DK.; Rodrigue, JR.; Taylor, WC., editors. Health promoting and health-compromising behaviors among minority adolescents. Washington, DC: American Psychological Association; 1997. p. 87-105.
- Rogers, RW. Cognitive and physiological processes in fear appeals and attitude change: A revised theory of protection motivation. In: Cacioppi, T.; Petty, RE., editors. *Social psychology*. New York: Guilford; 1983. p. 153-176.
- Rotheram-Borus MJ, Rebchook GM, Kelly JA, Adams J, Neumann MS. Bridging research and practice: Community-researcher partnerships for replicating effective interventions. *AIDS Education and Prevention* 2000;12(Suppl A):49–61. [PubMed: 11063069]
- Simpson DD. A conceptual framework for transferring research to practice. *Journal of Substance Abuse Treatment* 2002;22:171–182. [PubMed: 12072162]
- Stanton BF, Li X, Black M, Feigelman S, Ricardo I, Galbraith J, et al. Development of a culturally, theoretically and developmentally based survey instrument for assessing risk behaviors among inner city African-American early adolescents. *AIDS Education and Prevention* 1995;7:160–177. [PubMed: 7619645]
- Stanton BF, Li X, Ricardo I, Galbraith J, Feigelman S, Kaljee L. A randomized controlled effectiveness trial of an AIDS prevention program for low-income African-American youth. *Archives of Pediatrics and Adolescent Medicine* 1996;150:363–372. [PubMed: 8634730]
- Vincent ML, Paine-Andrews A, Fisher J, Devereaux R, Dolan HG, Harris KJ, et al. Replication of a community-based multicomponent teen pregnancy prevention model: Realities and challenges. *Family and Community Health* 2000;23:28–45.
- Visser MJ, Schoeman JB. Implementing a community intervention to reduce young people's risks for getting HIV: Unraveling the complexities. *Journal of Community Psychology* 2004;32:145–165.