



Published in final edited form as:

J Subst Abuse Treat. 2013 October ; 45(4): 363–369. doi:10.1016/j.jsat.2013.05.009.

Measuring fidelity to a culturally adapted HIV prevention intervention for men in substance abuse treatment

Mary Hatch-Maillette, Ph. D.^{1,2}, A. Kathleen Burlew, Ph.D.³, Sharriann Turnbull, B.S.⁴, Michael Robinson, B.S.⁵, and Donald A. Calsyn, Ph.D.^{1,2}

¹Alcohol and Drug Abuse Institute, University of Washington, Seattle, WA

²Department of Psychiatry and Behavioral Science, University of Washington School of Medicine, Seattle, WA

³Department of Psychology, University of Cincinnati, Cincinnati, OH

⁴Department of Psychology, Howard University, Washington, DC

⁵Department of Sociology, Wake Forest University, Winston-Salem, NC

Abstract

A fidelity measure was developed for use with Real Men Are Safe-Culturally Adapted (REMAS-CA), an HIV prevention intervention for ethnically diverse men in substance abuse treatment. The aims of this analysis were to: 1) assess the reliability of the Fidelity Rating and Skill Evaluation (FRASE); 2) measure improvement in therapist competence and adherence over time while delivering REMAS-CA; and 3) identify which modules of REMAS-CA were most difficult to deliver. Results showed that, 1) the FRASE was a reliable instrument; 2) therapists achieved adequate adherence and competence after training and demonstrated significant improvement over time in Global Empathy; and 3) Sessions 4 and 5 of REMAS-CA contained the most challenging modules for therapists to deliver. Recommendations for future REMAS-CA therapist trainings and fidelity monitoring are made.

Keywords

Training; fidelity; HIV-prevention; substance abuse treatment

1. Introduction

Real Men Are Safe-Culturally Adapted (REMAS-CA), an HIV prevention intervention for men in substance abuse treatment, was recently piloted and showed promising efficacy (Calsyn, Burlew, Hatch-Maillette et al., 2013). To ensure correct implementation of any clinical intervention, therapist adherence and competence must be systematically evaluated

© 2013 Elsevier Inc. All rights reserved.

Corresponding author: Mary Hatch-Maillette, Ph.D., University of Washington Alcohol & Drug Abuse Institute, 1107 NE 45th St., Ste. 120, Seattle, WA, 98105, United States, phone: 206-616-7730, fax: 206-543-2861, hatchm@uw.edu.

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

Findings were presented in part during a poster session at the Annual Scientific Meeting of the College on Problems of Drug Dependence, La Quinta, CA, June 11, 2012.

using a standardized tool. No such tool as yet exists for REMAS-CA or for its parent intervention, Real Men Are Safe (REMAS; Calsyn, Hatch-Maillette, Tross et al., 2009).

1.2 REMAS Is Effective

In a large-scale randomized clinical trial within the National Drug Abuse Treatment Clinical Trials Network, REMAS (Calsyn et al., 2009) was shown to be effective at reducing unprotected sex. Specifically, Calsyn and colleagues reported that from a group of 590 treatment-seeking men, those randomized to the 5-session REMAS HIV prevention intervention engaged in fewer unprotected vaginal and anal intercourse occasions in the 90 days prior to the 3- and 6-month post intervention follow-ups compared to those randomized to a 1-session HIV Education intervention. Also, compared to men who attended the HIV-Education session, those attending REMAS were less likely to have engaged in sex under the influence of alcohol or drugs during their most recent sexual encounter prior to the 3-month follow-up. In other words, REMAS appeared to be effective at reducing unprotected vaginal and anal sexual activity and sex under the influence. However, post-hoc analyses showed that for an important outcome measure, condom use with casual partners, a differential outcome based on ethnicity existed: Whites benefited more from REMAS than did African-Americans. And, although there were not enough Hispanics to be included in the post-hoc analyses, follow-up data showed that Hispanics had not increased their condom use either.

1.3 REMAS-CA Is Promising

In response to this differential effect, Calsyn and colleagues revised REMAS to be more culturally relevant to a diverse group of treatment-seeking men (Calsyn, Burlew, Hatch-Maillette et al., 2012). REMAS-CA differs from REMAS in the following ways: 1) REMAS-CA includes more focus on the impact of culture, social norms and upbringing on sexual behaviors and relationships. In fact, one full session is devoted to family, neighborhood, and subcultural factors that influence relationships with women. 2) In response to expert feedback, REMAS-CA sessions include more active learning and fewer didactic presentations by the facilitators. 3) Each session ends with a Talking Circle, an activity modeled after a Native American tradition of passing around an object with symbolic meaning to the group (e.g., feather, coin) to provide an opportunity for each person to share any personal reactions to the session without interruption. 4) Movie-clips are included in REMAS-CA to facilitate discussion on ways to initiate safe sex in real life situations. 5) The language and the names used in case studies are both revised to be more familiar to the target group.

In a pilot study of 95 men in two methadone maintenance and two outpatient psychosocial treatment programs in the United States, REMAS-CA proved to be more effective for African-American and Hispanic men compared to the original REMAS intervention: minority participants showed improved treatment session attendance and reduced levels of unprotected sex with casual partners (Calsyn et al., 2013). Given these positive findings, a valid tool to assess therapist adherence and competence seemed necessary for future use with the manual.

1.4 Therapist Adherence and Competence in Substance Abuse and HIV Prevention Trials

Webb and colleagues (2010) defined “therapist adherence” as the extent to which a therapist delivers an intervention as specified in the manual, and “therapist competence” as the level of skill with which s/he delivers the intervention. These constructs have been measured in a variety of ways in prior HIV prevention and substance abuse treatment studies. For example, in Project RESPECT, a randomized controlled trial comparing the efficacy of three models of HIV prevention counseling for high risk individuals, Kamb and colleagues (1996) used

multiple tools to monitor adherence and competence. These included post-training supervision and session review by supervisors and third-party monitors, and assessments of participants' reports of what happened during each session. Supervisors and the independent observers used a measure with a 1 ("Not achieved") to 5 ("Excelled") scale to indicate how well counselors delivered specific pieces of each intervention (e.g., "Assessed barriers to HIV risk reduction") as well as their counseling skills (e.g. "Demonstrated professionalism throughout session"). The NIMH Multisite HIV Prevention Trial similarly used a centralized training, counselor certification procedures, and audiotaped sessions to monitor fidelity. Monitoring occurred via assessments of the extent of topic coverage (Poor/Adequate/Excellent) and the presence (1= not at all, 5=maximum) of process skills such as reinforcement, goal setting, role playing, and problem solving (NIMH Multisite HIV Prevention Trial, 1997). Finally, the NIDA Collaborative Cocaine Study's therapist adherence and competence scale also measured the frequency and quality with which the counselor engaged in a particular intervention during each session using a 1 (low) to 7 (high) scale (Barber, Mercer, Krakauer & Calvo, 1996).

1.5 Changes in Adherence and Competence Over Time

Changes in therapist performance after training jeopardize the efficacy of an intervention. Even though a number of studies have been conducted on therapist improvement over time, the findings have been inconsistent. Crits-Cristoph, Siqueland, Chittams et al (1998) demonstrated that therapists delivering cognitive therapy *improved* their skill over time and across clients. However, both Baer et al., (2004) and Miller and Mount (2001) reported a *deterioration* in skills if not accompanied by supervision or monitoring. Still other researchers reported that counselors who demonstrate adequate competence in motivational interviewing (MI) *maintained* their skills over the subsequent 24 weeks without additional training (Martino, Canning-Ball, Carroll, & Rounsaville, 2011). The inconsistent findings across protocols suggest a need to evaluate each protocol on its potential for therapist improvement, maintenance, or deterioration of skills over time after training.

Pertaining to this question of counselor skill development, Calsyn and colleagues conducted a series of post-intervention focus groups, one with participants and one with counselors, at each of the REMAS-CA pilot study sites. Investigators asked questions about what parts of the REMAS-CA intervention went well, what did not, and what was difficult for participants to understand or engage with. A key question was what session(s) and/or module(s) was most difficult to present to participants. Similar discussion topics were also broached on the monthly supervision calls with counselors while the study was running. In both settings, counselors consistently reported that they struggled with modules in Sessions 4 and 5 that pertained to how culture had influenced participants' views on gender roles, power, sexuality, and responsibility for sexual safety. Therefore, the authors sought to determine whether this anecdotal evidence regarding what were the more challenging modules in REMAS-CA was supported by objective measures of counselor competence, and whether their competence and skill changed over time post-training.

1.6 Study Aims

The purpose of this study was threefold: 1) to develop an adherence scale for REMAS-CA that can be paired with the manual and be a resource for treatment settings in their implementation; 2) to measure improvement in counselors' intervention delivery skills over time across cohorts; and 3) to identify which modules of REMAS-CA were most difficult to deliver, so as to guide future training and supervision resources appropriately.

2. Methods

2.1 Real Men Are Safe-Culturally Adapted (REMAS-CA)

REMAS-CA (Calsyn et al, 2012) is a five-Session HIV risk reduction intervention designed for men in substance abuse treatment. It is a culturally adapted, and recently piloted, version of Real Men Are Safe (Calsyn et al., 2009), an evidence-based intervention listed on HIV/AIDS Prevention Research Synthesis Project of the Center for Disease Control and Prevention (CDC, 2012). Two male counselors, at least one of whom was a person of color, delivered the intervention via 90-minute group sessions. Groups typically occurred over a three week period. Although a cohort of 8–10 men was considered to be ideal, it was not uncommon for groups to be smaller due to attendance barriers. The five sessions of REMAS-CA use a combination of didactic, skill-building, and motivational methods to deliver content around HIV/safe sex education, relationship/communication skills, negotiating safe sex, the overlap of sex and drug use, and the ways in which the men's self-described culture (e.g., ethnic, social, religious) have influenced their attitudes toward sex and drug risk. Session titles are: Session 1, HIV/AIDS Update/Identifying Risks; Session 2, HIV/AIDS Update/Planning & Prevention; Session 3, Sex Without Drugs: Can It Happen/Is It Pleasurable?; Session 4, Intimate Relationships: Understanding the Origins of Attitudes, Beliefs, Hopes; and Session 5, Beyond the Pick-Up Line— Communicating About Sex.

2.2 Performance Sites

The data for this study were collected during a pilot study of REMAS-CA at four substance abuse Community Treatment Programs (CTPs) within the NIDA Clinical Trials Network (CTN). Two sites provided opioid agonist therapy (Matrix Institute, Los Angeles, CA and Hartford Dispensary, Hartford, CT). The other two sites were psychosocial treatment programs offering a wide variety of treatment options with most participants being recruited from their intensive outpatient treatment programs (Lexington-Richmond Alcohol and Drug Abuse Council, Columbia, SC and The Life Link, Santa Fe, NM).

2.3 Study Counselors and Supervision

Each CTP identified two clinical staff per site to serve as REMAS-CA counselors. Of these eight counselors, all but one was recruited from the existing counselor pool at the agencies. One counselor was hired specifically for the study because the agency did not have enough existing staff time available. All were experienced addiction treatment providers; most had chemical dependency certifications and one had a doctorate in clinical psychology. These interventionists attended a centralized two day training provided by Drs. Calsyn, Hatch-Maillette, and Burlew. Counselors were certified to deliver REMAS-CA after submitting two rounds of audio-recorded role play groups (one full session and two additional modules).

Counselors received supervision during their certification process and throughout the duration of the pilot study from the study trainers/investigators, Drs. Calsyn and Hatch-Maillette. For certification, the trainers reviewed each audio recording the counselors made, provided written feedback, and used the standardized rating scale from the original REMAS protocol training to rate therapist competence and adherence. If the counselors did not meet minimum adherence and competence criteria (defined as a score of 3 out of 5 on both domains), written comments along with their FRASE ratings, and sometimes additional verbal feedback, was provided to the counselor, who was instructed to practice further with additional role plays and then submit another recording for evaluation.

Ongoing supervision during the course of the pilot study was provided by Drs. Calsyn and Hatch-Maillette, via regularly scheduled monthly conference calls. These calls provided a

format in which counselors could ask questions, talk with their counterparts at other sites about their experiences delivering REMAS-CA, and receive any general feedback that the trainers had noted in their fidelity reviews of submitted audio-recorded sessions.

2.4 Fidelity Monitoring

The adherence and competence measure used in this study, Fidelity Rating and Skill Evaluation (FRASE; see Appendix 1) for REMAS-CA, was developed from a similar measure by Gudyish and colleagues (Campbell, Manuel, Turcotte Manser et al., 2013) and is consistent with recommendations from Waltz, Addis, Koerner and Jacobson's (1993) review of strategies for evaluating treatment integrity. They recommended that first, therapeutic competence should be related directly to a treatment manual that explains how the therapy should be implemented. Second, "competence" or skill with an intervention should not be confused with "adherence" or "frequency" of certain behaviors or intervention components. A final recommendation was that adherence measures include items that are unique and essential to the treatment, behaviors that are essential but not unique to it, behaviors that are acceptable but not necessary, and behaviors that are contraindicated (proscribed).

The FRASE measure followed these recommendations by assessing several aspects of counselors' performance: Extensiveness (how thoroughly they delivered each module), Competence (how skillfully they delivered each module), Proscribed Behaviors (avoidance of therapist behaviors that deviated from the spirit of psychoeducational group therapy intervention), Co-Therapy (skill in co-leading a group), Global Empathy (overall rating of therapist empathy), and Global Skill (overall rating of therapist-pair skill). Depending on the session, Extensiveness and Competence domains contained 10–14 items total, while Proscribed Behaviors, Co-Therapy, Global Empathy, and Global Skill each contained 3, 4, 1 and 1 items, respectively. Items were rated on a 1 (Not At All) to 5 (Extensively) scale. Examples included, "*To what extent did counselors present the Condom Demonstration module?*" (Extensiveness), "*How well did the counselors present the Condom Demonstration module?*" (Competence), "*To what extent did the counselors present didactic material in an overly structured, non-interactive manner?*" (Proscribed Behavior), "*How well did the co-therapy team handle time management?*" (Co-Therapy), "*Overall, how well did the counselors understand and/or attempt to grasp the clients' perspectives?*" (Global Empathy), and "*Overall, how well did the counselors conduct this specific group Session?*" (Global Skill).

2.5 Procedures

The inter-rater reliability coding team of four consisted of the lead- and one co-investigator, and two undergraduates. The team reviewed group session recordings for the domains outlined above. For each cohort of participants, there were 20 possible intervention sessions to be rated (5 sessions x 4 sites).

2.5.1 Undergraduate Rater Training—Two undergraduate raters were trained for this analysis by Drs. Calsyn and Hatch-Maillette. Training consisted of orientation to REMAS and REMAS-CA, a review of concepts pertaining to fidelity monitoring (e.g. counselor adherence/competence, rationale for monitoring in multi-site protocols, use of a coding manual and rating forms), and a graduated process of audio-recording review until their ratings agreed enough with the expert (investigators) ratings to be considered independent.

Cohort 2 at each site was used as the fidelity coding training cohort for all four raters. Cohort 2 was chosen because this study intended to compare ratings from counselors' first (Cohort 1) and last (Cohort 3 or 4, depending on site) cohorts to assess change over time.

Since raters needed “training cases” to establish their inter-rater reliability, a cohort whose ratings would not be included in the final analysis had to be selected.

Although there were a total of 20 Cohort 2 sessions delivered across sites, only fifteen session recordings were available (five were unavailable due to malfunctions with recording equipment). From the 15 available, five sessions were randomly chosen to be rated. One session was rated by all four raters simultaneously, as a way to orient to using the coding manual. The remaining four sessions were independently rated by each of the four raters. After reviewing a session, the raters met to discuss their rationale for their ratings with the goal of achieving a consensus in ratings. These discussions at times led to modest revisions in the fidelity rating manual as a means of bolstering consistency across raters and improving precision. For the remaining ten Cohort 2 recordings, the undergraduate raters coded all sessions while each investigator coded 5. The four raters again met to discuss discrepancies and further clarify and refine the fidelity rating manual procedures and definitions. Session coding took approximately 90–100 minutes as it required listening to the entire 90 minute REMAS-CA session. Consensus meetings, held during the undergraduate rater training, were scheduled for 120 minutes to allow sufficient time to listen to a session together, code, and discuss discrepancies.

2.5.2 Adherence and Competence Ratings—For Cohort 1 and the final Cohort (either 3 or 4, depending on the agency’s number of cohorts) there were 33 out of a possible 40 session recordings available to be rated. Seven sessions were not available to be rated because of technology failures such as dead batteries, poor volume control, not having started the recorder, recording over a previously existing audio file, or loss of audio files during electronic file transfer. The undergraduates rated all 33 of these recordings independently. The undergraduates rated blindly and were unaware of our hypothesis that sessions 4 and 5 would be more difficult for counselors to deliver.

2.6 Data Analysis

2.6.1 Inter-rater reliability—To assess inter-rater reliability among raters, the intra-class correlation coefficient (ICC) was calculated using a mixed effects model with absolute agreement via the Reliability Program from SPSS. The ICC was calculated for all four raters, 2 sets of three raters and for the two raters who rated all sessions.

2.6.2 Counselor improvement/Session difficulty—Ratings were summed across adherence items (5 to 7 per session, one for each module) so that for each session a mean extensiveness and a mean competence value were obtained. Similarly, the Proscribed Behavior and Co-Therapy items were averaged. Thus for each session there were four averaged ratings as described above, plus the ratings for global empathy and global skill. To determine if counselors improved in their delivery of the intervention over time, these six scores were compared between first and last cohort delivered during the study utilizing paired t-tests. To determine which modules were easier or more difficult to deliver, extensiveness and competence means from the final cohort were calculated for each module. The choice to use the final cohort scores to determine session difficulty was based on the assumption that after running two to three cohorts, counselors would have had a chance to “work the kinks out.” Observations of counselors during training and their anecdotal reports on subsequent counselor supervision conference calls, revealed that overall familiarity with the material and with following a manualized psychoeducational group format were skills that counselors needed to hone, in addition to mastering the delivery of specific REMAS-CA content. Therefore, we chose the final cohort as the “best” example of counselors’ work because they had had the most practice, and lower extensiveness or competence scores might reveal aspects of true difficulty with the material delivery.

3. Results

3.1 Inter-rater reliability

Results showed that the FRASE could be used reliably by both doctoral and undergraduate level raters. The single measure ICC for all four raters (105 items) on the FRASE was 0.845, which shows “excellent” inter-rater reliability (Cicchetti, 1994). The single measure ICCs for three raters (210 items) were 0.618 and 0.696 respectively, and the single measure ICCs for two independent undergraduate raters (1006 items) was 0.734. According to Cicchetti (1994), these ICCs fall within the range of “good” inter-rater reliability. The lowest ICC of 0.618 did not appear to be due to any one rater, but rather to those sessions that were coded by the group of three versus two raters. Our results confirmed that the training in coding procedures using the Cohort 2 recordings had achieved a level of consistency across raters even when the assistance and guidance from the investigators was absent.

3.2 Counselor improvement across cohorts

The counselors demonstrated improvement on the global empathy measure but not the other scales (Table 1).

3.3 Session difficulty

Counselors received higher extensiveness ratings versus skill ratings for all modules, indicating that covering the material did not equate to covering it well. The scores on the Talking Circle were consistently lower than the scores on the other session activities. In general, it appears that modules that are primarily informational -- such as many of those in sessions 1 and 2 -- are easiest to deliver skillfully. More difficult are those requiring an exploration of current or past interpersonal relationships, changing social norms, or communication skills, such as those modules in Sessions 4 and 5 (Table 2).

4. Discussion

A recent study of the implementation of evidence based HIV prevention interventions revealed that many therapists either modified core characteristics or reinvented the intervention by adding new material and even replacing key characteristics (Veniegas, Kao, & Rosales, 2009). Fidelity monitoring enables researchers to evaluate the effectiveness of HIV interventions more accurately by assessing the extent to which the interventions are implemented as intended. The availability of the FRASE for monitoring the fidelity of REMAS-CA, a promising intervention for reducing HIV risk behaviors in ethnically diverse groups of substance abusing men, will facilitate future research on REMAS-CA. Moreover, the FRASE has both the potential for assessing change in counselor performance over time and the capacity to identify the modules most difficult to implement. The findings of this study, although preliminary due to the size of this pilot study sample, can guide REMAS-CA researchers and others when training counselors to implement the intervention.

4.1 Training procedures for enhancing treatment fidelity

The level of therapist fidelity obtained, coupled with the absence of drift, suggest that our training procedures were adequate. Our procedures were similar to those recommended by others. As suggested by Borrelli (2011), we included a standardized training plan to ensure that therapists are trained to criterion. Our training plan included both the indirect (didactic presentations) and direct (e.g., opportunities for practice such as role-plays and other forms of rehearsal, demonstrations, and feedback) training activities proposed by Pereplechikova et al (2005). Finally, our training plan included the ongoing monitoring and supervision that Pereplechikova et al., argue are essential to therapist fidelity.

4.2 The FRASE Can Be Used to Assess Fidelity to REMAS-CA

The FRASE includes ratings of counselor adherence (i.e., extensiveness) and competence (i.e., competence, proscribed behaviors, co-therapy, and global empathy). Overall, the moderate to high inter-rater reliabilities in this study suggested that agreement among raters is achievable, making the FRASE an implementable and promising tool for assessing both adherence and competence in implementing REMAS-CA.

4.3 Therapist Improvement

The findings on therapist improvement in the delivery of the intervention were mixed. Interestingly, therapist ratings on global empathy were higher on the last than the first cohort. One explanation is that therapists' empathy grew as the men shared intimate details of the ways in which their life experiences, cultural values, and social norms shaped their sexual behaviors and interactions with women. We wonder whether therapists' increased understanding of this target group manifested in higher global empathy scores. Another possibility is that, as the counselors grew more comfortable with the delivery of the intervention activities, their empathy increased because they were able to focus more on the nuances and the stories of the participants rather than the mechanics of intervention delivery.

Therapist ratings did not change on the other performance indicators. The high initial ratings on the other indicators (e.g., Extensiveness, Competence, Proscribed Behaviors, and Co-Therapy) may provide a possible explanation of why therapists only demonstrated improvement on the global empathy indicator. The high ratings on the other indicators may have created a ceiling effect and, consequently, limited the likelihood of detecting improvement on those measures.

As mentioned earlier, Martino et al., (2011) reported that their therapists maintained but did not increase in performance during their 24 week clinical trial. Our findings on the Extensiveness, Competence, Proscribed Behaviors, and Co-Therapy skills are similar to the Martino et al findings. However, our findings on global empathy are similar to the improvement in therapist performance that Crits-Christoph et al., (1998) reported. Several important differences between the counselor training in the Crit-Christops et al study and our REMAS study may explain why our therapist improvement was more limited than the improvement reported by Crit-Christoph et al. First, the intensity of the ongoing training and supervision in the Crits-Christoph et al study exceeded what we were able to provide in the REMAS-CA study. Even though we scheduled monthly counselor supervision conference calls in REMAS-CA, not all counselor were able to attend regularly due to clinical and scheduling demands. Therefore, the supervision occurred on a voluntary rather than a mandatory basis, and this likely lessened the benefit that post-training supervision might offer. Second, although we rated the therapists on adherence and competence, the Crits-Christoph team not only rated their therapists but shared the ratings with therapists in ongoing supervision sessions. One conclusion is that training prior to implementing the intervention may be adequate to prepare therapists to grow in therapist empathy. However, ongoing training and a supervision that informs therapists about their performance across sessions may increase their competence in other skills over time.

The amount of training and supervision provided by the Crits-Christoph team is labor-intensive and expensive. Fortunately, the aforementioned study of MI therapists by Martino et al., (2011) suggests that not all therapists may require additional ongoing training. Specifically, in that study, therapists were classified as either competent or in need of further training at the end of the initial training. Additional training was provided only to the latter group, who then showed improvement over time in adherence and competence. The therapists who did not require further training maintained their adherence and competence

skills over time. This finding suggests that researchers and clinical supervisors may reduce the cost in time and money by identifying and providing additional training only to those therapists whose initial competence ratings suggest they would benefit from further training.

Instruments such as the FRASE can also be used to detect a decrement in counselor skills (e.g., drift) in implementing the intervention curriculum. It is notable that very little drift was evident in the FRASE ratings. This finding supports the adequacy of the training used to prepare the counselors in our recent research study (Calsyn et al, 2013). The presence of considerable drift may have argued for more intensive training at the beginning, more monitoring of therapist performance perhaps with ongoing feedback and supervision, or even booster sessions.

4.4 Session Difficulty

The mean ratings suggest that certain sessions are more difficult for counselors to implement than others. Based on the ratings, the counselors experienced more difficulty implementing Sessions 4 and 5 than the earlier sessions. This finding was consistent with anecdotal reports from counselors and participants throughout the study. Session 4 explores the cultural origins of the participants' own beliefs about power, love, and communication in relationships, how to change their social/cultural norms as they relate to those topics, and how to communicate assertively. The participants described this session as the most emotionally-charged of the five sessions. Both Session 4 and 5 provide training on acquiring a specific skill: communicating assertively about sexual risk behavior. None of the other modules are as emotionally charged or challenge the participants to practice new skills that may be difficult for them to learn. In fact, Sessions 1, 2 and 3 involve more didactic and less interactive presentation of material, thereby suggesting that therapists have an easier time with the former. The lower ratings on the final two sessions suggest that trainers might devote more time to preparing counselors to implement Sessions 4 and 5 compared to Sessions 1, 2 and 3. Indeed, as Perepletchikova and Kazdin (2005) suggest, procedures such as, role plays, the use of videotapes to model the implementation of the activity, rehearsals, ongoing monitoring, and feedback on adherence, may enhance therapist skill in implementing sessions 4 and 5.

The Talking Circles were rated lower on skill in every module than any of the other activities in the module. Although we have no empirical data to explain this pattern, one possible explanation for the low skill scores is that the Talking Circles, since they occurred at the end of sessions, may have been rushed when the therapists were running behind. Another possibility is that the purpose and format of the Talking Circles requires more attention.

4.5 Limitations

This study has several limitations. First, REMAS-CA was conducted in clinics motivated to implement the intervention in a research project. The counselors, who had all volunteered to engage in intervention development work, might have felt more ownership and commitment to implement the intervention successfully than counselors in more neutral or less invested settings. Iverson and colleagues (2008) has suggested that adherence may increase with provider buy-in. Therefore more research may be required to determine whether the training dose and the format described in this study are appropriate for a diverse group of clinics who are adopting REMAS-CA for the first time.

Second, though we were not aware of any tendencies for inflating the initial ratings of the therapists, using members of the research team as raters may have posed a potential conflict of interest. Specifically, research team members who are concerned about meeting study

timelines may inadvertently be more forgiving in their initial FRASE ratings so as to avoid costly delays in study implementation. While higher ratings at the beginning may allow the therapists to be certified and begin the intervention earlier, ultimately they may restrict the range on the FRASE to the point that the therapists have less room to show improvement. For example, counselors post-training may have performed in the “Fair” (2) to “Adequate” (3) range, but because certification required a score of 3 or above, raters may have been more inclined to score counselors with a 3 than a 2. Thus, raters may not have given enough sub-competent scores early on, such that improvement was not captured later. For this reason, it may be more appropriate, and ultimately more helpful to the research team, to use external raters in future studies.

Third, the lack of counselor improvement over time may have been related to a ceiling effect. As noted in Rounsaville, Chevron, Weissman et al. (1986) and Crits-Christoph et al. (1998), the counselors in the current study were selected with care and likely were “the best suited” in their clinics to function as study counselors and deliver a manualized intervention. Therefore their performance may not have had as much room to show improvement over time because it was fairly strong from the outset.

Fourth, the results of the current study should be considered preliminary, as they are based on a pilot study that is smaller and less rigorous than a large-scale clinical trial.

4.6 Strengths

This study has several strengths. First, a national team of eight counselors implemented the intervention in four different community treatment programs. This geographic and treatment program variability allowed for reasonable generalizability to the target population. Second, the same research team that developed both REMAS and REMAS-CA conducted the training and the fidelity ratings, providing consistency across studies. Finally, the sessions were all audio-recorded so that the raters could review them more than once if necessary to bolster confidence in their ratings.

4.7 Future Directions

Several areas deserve more attention in future study. First, the emphasis in adapting REMAS was on modifying the intervention activities to be more culturally appropriate. Little attention to date has been devoted to ensuring that the counselors deliver the intervention activities in a culturally competent manner. (e.g. use of familiar language, use of examples that demonstrate an understanding of the social context in which the participants live). Similarly, the FRASE at this point only evaluates the accurate implementation of the curriculum. Future modifications to the FRASE might include some evaluation of the extent to which the revised curriculum is being delivered in a culturally competent manner. Of course, the raters can only accurately assess cultural competence in the delivery if the researchers can identify some specific behaviors related to the culturally competent delivery of a protocol and then train raters to assess these behaviors in a consistent manner. Therefore, the raters may require additional training on the criteria to be used to assess the delivery of the intervention.

The feasibility of the 90 minutes required to review a session audio-recording tape to implement the FRASE rating system is unclear. The use of the FRASE in future research in which the investigators may have grant funding to hire raters seems quite feasible. However, using the FRASE to monitor therapist performance in non-research settings in which clinics may already be straining to meet their productivity targets may be more of a challenge. One possibility would be to instruct supervisors to review fewer than all 5 sessions, for the purpose of “spot-checking” sessions.

Establishing the utility of the FRASE opens up the possibility of addressing several important questions. One question is the extent to which treatment fidelity is associated with better outcomes. However, an equally important question is whether therapist improvement in fidelity over time is associated with better outcomes over time. Even though the improvement in global empathy in this study between the first and third cohort was statistically significant, the actual differences in empathy were modest. Therefore, it remains an open question whether the significant improvement in global empathy was sufficient to improve clinical outcomes.

4.8 Summary

This study, the first to develop a measure of counselor fidelity for REMAS-CA, makes several important contributions for research and clinical practice. First, the ICC demonstrates that the FRASE is a viable tool for researchers and clinicians interested in assessing fidelity in implementing REMAS-CA. Second, the findings offer some guidelines for increasing counselor competence in implementing REMAS-CA for research or clinical practice. Specifically, more effort should be spent on training the counselors to implement the emotionally charged discussion and the specific skill building present in Sessions 4 and 5.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The research was supported by NIDA grant 1RC1DA028245-01 (D. Calsyn, PI), and NIDA grant 3K23DA025678-02S1 (B. Hartzler, PI).

The investigators wish to thank the eight counselors, four research coordinators, four site-PI's, and the patients and staff from the four community treatment programs involved in the trial for their assistance in conducting project.

The authors would like to acknowledge the significant contribution that Dr. Don Calsyn, Lead Investigator of the REMAS studies, made to this manuscript prior to his unexpected and sudden death in February 2013.

References

- Barber JP, Mercer D, Krakauer I, Calvo N. Development of an adherence/competence rating scale for individual drug counseling. *Drug and Alcohol Dependence*. 1996; 43:125–132. [PubMed: 9023068]
- Borrelli B. The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *Journal of Public Health Dentistry*. 2011; 71:S52–S63.
- Calsyn D, Burlew AK, Hatch-Maillette MA, Wilson J, Beadnell B, Wright L. Real Men Are Safe-Culturally Adapted: Utilizing the Delphi Process to Revise Real Men Are Safe for an Ethnically Diverse Group of Men in Substance Abuse Treatment. *AIDS Education and Prevention*. 2012; 24:118–132.
- Calsyn DA, Burlew AK, Hatch-Maillette MA, Beadnell B, Wright L, Wilson J. “Real Men Are Safe-Culturally Adapted,” an HIV-Prevention Intervention for Ethnically Diverse Men in Substance Abuse Treatment: Pilot Study Findings. *American Journal of Public Health*. 2013; 103:896–902. [PubMed: 23488494]
- Calsyn D, Hatch-Maillette M, Tross S, Doyle S, Crits-Christoph P, Song Y, Harrer J, Lalos G, Berns S. Motivational and skills training HIV/sexually transmitted infection sexual risk reduction groups for men. *Journal of Substance Abuse Treatment*. 2009; 37:138–150. [PubMed: 19150206]
- Campbell BK, Manuel JK, Turcotte Manser S, Peavy KM, Stelmokas J, McCarty D, Guydish J. Assessing Fidelity of Treatment Delivery in Group and Individual 12-Step Facilitation. *Journal of Substance Abuse Treatment*. 2013; 44:169–176. [PubMed: 22944595]

- Center for Disease Control and Prevention. [Accessed September 20, 2012] Compendium of Evidence-Based HIV Behavioral Interventions: HIV/AIDS Prevention Research Synthesis Project. 2012. <http://www.cdc.gov/hiv/topics/research/prs/index.htm>
- Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*. 1994; 6:284–290.
- Crits-Christoph P, Siqueland L, Chittams J, Barber J, Beck AT, Frank A, Liese B, Luborsky L, Mark D, Mercer D, Simon Onken L, Najavits LM, Thase ME, Woody G. Training in cognitive, supportive-expressive, and drug counseling therapies for cocaine dependence. *Journal of Consulting and Clinical Psychology*. 1998; 66:484–492. [PubMed: 9642886]
- Iverson EF, Balasuriya D, Garcia GP, Sheng M, Richardson JL, Stoyanoff S, King JB. The challenges of assessing fidelity to physician-driven HIV prevention interventions: Lessons learned implementing Partnerships for Health in a Los Angeles HIV clinic. *AIDS and Behavior*. 2008; 12:978–988. [PubMed: 18427973]
- Kamb MK, Dillon BA, Fishbein M, Willis KL. The Project RESPECT Study Group. Quality assurance of HIV prevention counseling in a multi-center randomized controlled trial. *Public Health Reports*. 1996; 111(S1):99–107. [PubMed: 8862164]
- Martino S, Canning-Ball M, Carroll KM, Rounsaville R. A criterion-based stepwise approach for training counselors in motivational interviewing. *Journal of Substance Abuse Treatment*. 2011; 40:357–365. [PubMed: 21316183]
- NIMH Multisite HIV Prevention Trial. *AIDS*. 1997; 11(Supplement 2):S49–S53. [PubMed: 9475711]
- Perepletchikova F, Kazdin AE. Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical Psychology: Science and Practice*. 2005; 12(4) 365?383.
- Rounsaville BJ, Chevron ES, Weissman MM, Prusoff BA, Frank E. Training therapists to perform interpersonal psychotherapy in clinical trials. *Comprehensive Psychiatry*. 1986; 27:364–371. [PubMed: 3524986]
- Veniegas RC, Kao UH, Rosales R. Adapting HIV prevention evidence-based interventions in practice: an interview study. *Implementation Science*. 2009;4. [PubMed: 19203373]
- Walters ST, Matson SA, Baer JS, Ziedonis DM. Effectiveness of workshop training for psychosocial addiction treatments: A systematic review. *Journal of Substance Abuse Treatment*. 2005; 29:283–293. [PubMed: 16311181]
- Waltz J, Addis ME, Koerner K, Jacobson N. Testing the integrity of a psychotherapy protocol: Assessment of adherence and competence. *Journal of Consulting and Clinical Psychology*. 1993; 61:620–630. [PubMed: 8370857]
- Webb CA, DeRubeis RJ, Barber JP. Therapist adherence/competence and treatment outcome: A meta-analytic review. *Journal of Consulting and Clinical Psychology*. 2010; 78:200–211. [PubMed: 20350031]

Table 1

Change in counselor adherence ratings from first to last cohort

	First Cohort	Last Cohort		
Rating Variable	Mean (sd)	Mean (sd)	Paired <i>t</i>	<i>p</i>
Extensiveness of Module Coverage	4.66 (0.32)	4.54 (0.48)	0.76	0.46
Skill of Module Delivery	4.09 (0.39)	3.92 (0.43)	1.23	0.24
Use of Proscribed Behaviors	1.41 (0.43)	1.37 (0.25)	0.30	0.77
Quality of Co-Therapy	3.77 (0.52)	4.01 (0.82)	1.12	0.28
Global Empathy	4.44 (0.77)	4.76 (0.53)	2.28	0.04
Global Skill	4.12 (0.72)	4.03 (0.60)	0.43	0.68

Table 2
FRASE extensiveness and skill ratings for REMAS-CA modules delivered to last study cohort

Session 1 Modules	Extensiveness	SD	Range	Skill	SD	Range
Introduction, Review and Check-in	5.00	0.00	5.0–5.0	4.00	0.71	3.5–5.0
Risky Behaviors Exercise	5.00	0.00	5.0–5.0	4.38	0.95	3.0–5.0
HIV/AIDS Update	4.88	0.25	4.5–5.0	4.00	0.41	3.5–4.5
HIV Risky Behaviors Exercise-Revisited	4.67	0.29	4.5–5.0	4.50	0.00	4.5–4.5
Talking Circle	4.33	1.15	3.0–5.0	2.50	0.00	2.5–2.5
<i>Means for Session 1</i>	<i>4.78</i>			<i>3.88</i>		
Session 2 Modules						
Introduction, Review and Check-in	4.75	0.50	4.0–5.0	4.25	0.96	3.0–5.0
Review Drug Use Practices and Safe Sex Hierarchies	4.75	0.29	4.5–5.0	3.75	0.29	3.5–4.0
Conduct Condom Demonstration	4.50	0.41	4.0–5.0	4.38	0.25	4.0–4.5
Male Condom Practice	4.67	0.29	4.5–5.0	4.83	0.29	4.5–5.0
Female Condom Demonstration	4.50	0.41	4.0–5.0	4.38	0.48	4.0–5.0
Barriers to Condom Use: Brainstorming Solutions	4.63	0.25	4.5–5.0	4.38	0.25	4.0–4.5
Talking Circle	3.50	1.73	1.0–5.0	2.50	1.29	1.0–4.0
<i>Means for Session 2</i>	<i>4.47</i>			<i>4.07</i>		
Session 3 Modules						
Introduction, Review and Check-in	4.13	1.03	3.0–5.0	3.38	0.75	2.5–4.0
Building Skills for Making and Communicating Safer Sex Decisions: Movie Clips	4.50	0.71	3.5–5.0	3.88	1.31	2.0–5.0
Experience with Sex and Drugs Brainstorming	4.75	0.29	4.5–5.0	4.13	0.75	3.5–5.0
Enhancing Sex without Drugs Brainstorming	4.75	0.29	4.5–5.0	4.38	0.48	4.0–5.0

Session 1 Modules	Extensiveness	SD	Range	Skill	SD	Range
Talking Circle	3.62	1.75	1.0-4.5	3.00	1.58	1.0-4.5
<i>Means for Session 3</i>	4.35			3.75		
Session 4 Modules						
Introduction, Review and Check-in	4.88	0.25	4.5-5.0	4.38	0.48	4.0-5.0
“Who’s Got the Power? Who’s Showin the Love?”	4.63	0.48	4.0-5.0	3.88	0.25	3.5-4.0
Cultural Values and Intimate Relationships	4.33	0.76	3.5-5.0	3.67	0.29	3.5-4.0
Ideal Man/Ideal Woman	4.50	0.50	4.0-5.0	4.00	0.50	3.5-4.5
Changing Social Norms	3.83	1.61	2.0-5.0	3.67	1.53	2.0-5.0
Talking Circle	2.83	1.61	1.0-4.0	1.83	1.04	1.0-3.0
<i>Means for Session 4</i>	4.17			3.57		
Session 5 Modules						
Introduction, Review and Check-in	4.17	1.44	2.5-5.0	3.67	1.04	2.5-4.5
Communicating About Safe Sex: TALK Tools	4.83	0.29	4.5-5.0	4.33	0.29	4.0-4.5
Practice TALK Tools: Responding to Excuses for Not Using Condoms	4.83	0.29	4.5-5.0	4.17	0.29	4.0-4.5
Practice TALK tools: Responding to Come-Ons for Sex Under the Influence	3.17	2.02	1.0-5.0	3.83	1.61	1.0-4.0
TALK Role Plays	3.33	2.08	1.0-5.0	3.17	2.02	1.0-5.0
Personalizing Commitment to Sexual Safety	4.83	0.29	4.5-5.0	4.00	0.50	3.5-4.5
Talking Circle	4.83	0.29	4.5-5.0	3.00	0.87	2.0-3.5
<i>Means for Session 5</i>	4.28			3.74		