

Understanding Leisure-Related Program Effects by Using Process Data in the HealthWise South Africa Project

Linda L. Caldwell
Anita S. Younker
Lisa Wegner
Megan E. Patrick
Tania Vergnani
Edward A. Smith
Alan J. Flisher

EXECUTIVE SUMMARY: As the push for evidence-based programming gathers momentum, many human services programs and interventions are under increased scrutiny to justify their effectiveness across different conditions and populations. Government agencies and the public want to be assured that their resources are being put to good use on programs that are effective and efficient. Thus, programs are increasingly based on theory and evaluated through randomized control trials using longitudinal data. Despite this progress, hypothesized outcomes are often not detected and/or their effect sizes are small. Moreover, findings may go against intuition or “gut feelings” on the part of project staff. Given the need to understand how program implementation issues relate to outcomes, this study focuses on whether process measures that focus on program implementation and fidelity can shed light on associated outcomes. In particular, we linked the process evaluation of the HealthWise motivation lesson with outcomes across four waves of data collection. We hypothesized that HealthWise would increase learners’ intrinsic and identified forms of motivation, and decrease amotivation and extrinsic motivation. We did not hypothesize a direction of effects on introjected motivation due to its conceptual ambiguity. Data came from youth in four intervention schools ($n = 902$, 41.1%) and five control schools ($n = 1291$, 58.9%) who were participating in a multi-cohort, longitudinal study. The schools were in a township near Cape Town, South Africa. For each cohort, baseline data are collected on learners as they begin grade 8. We currently have four waves of data collected on the first cohort, which is the focus of this paper. The mean age of the sample at wave 3 was 15.0 years ($SD = .86$) and 51% of students were female. Results suggested that there was evidence of an overall program effect of the curriculum on amotivation regardless of fidelity of implementation. Compared to the control schools, all treatment school learners reported lower levels of amotivation in wave 4 compared to wave 3, as hypothesized. Using process evaluation data to monitor implementation fi-

delity, however, we also conclude that the school with better trained teachers who also reported higher levels of program fidelity had better outcomes than the other schools. We discuss the implications of linking process data with outcome data and the associated methodological challenges in linking these data.

Keywords: Adolescents, implementation fidelity, leisure motivation, prevention, process data.

Authors: Caldwell is with the Department of Recreation, Park, and Tourism Management at The Pennsylvania State University, 801 Donald A. Ford Building, University Park, PA, 16802. Email: lindac@psu.edu. Younker, Patrick, and Smith are with the same institution. Wegner is with the University of the Western Cape; Vergnani is also with the University of the Western Cape; and Flisher is with both the University of Cape Town and the University of Bergen. Support for this research was provided by the National Institute on Drug Abuse (NIDA; DA017629, DA10075, DA017491). The authors would like to thank the reviewers for comments that significantly improved the paper.

As the push for evidence-based programming gathers momentum, many human services programs and interventions are under increased scrutiny to justify their effectiveness across different conditions and populations. Government agencies and the public want to be assured that their resources are being put to good use on programs that are efficient and effective (Guskey, 2000). Thus, programs are increasingly based on theory and evaluated through randomized control trials using longitudinal data. Despite this progress, hypothesized outcomes are often not detected and/or their effect sizes are small (Gingiss, Roberts-Gray, & Boerm, 2006). Moreover, findings may go against intuition or “gut feelings” on the part of project staff. A related issue is the process of taking an evidence-based program that is successful under tightly controlled conditions (e.g., an efficacy trial where researchers have more control over program implementation) to the field, where there is much less control over context and implementation. The purpose of this study is to address these issues by examining associations between process implementation fidelity and program outcomes. Specifically, we focus on how process implementation fidelity data from a lesson on leisure motivation can inform short-term, motivation-related outcomes. Data for the investigation are derived from an evaluation trial of a substance use and HIV/AIDS prevention intervention for a sample of South African youth.

Program Implementation and Fidelity

Lack of support for hypothesized findings in a given evaluation might not indicate a poor program, but rather may stem in part from poor implementation or lack of fidelity to the intervention as planned (Bellg, et al., 2005; Dusenbury, Brannigan, Falco, & Hansen, 2003). Often overlooked until recently, intervention implementation fidelity

data (i.e., process data) are being collected and used to understand intended program outcomes, as well as adoption, diffusion, and/or the cultural adaptation of empirically validated programs (c.f., Dusenbury, Brannigan, Falco, & Hansen, 2003; Elliott & Mihalic, 2004; Gingiss, Roberts-Gray, & Boerm, 2006). These data help researchers and program providers understand when a Type III error has occurred—that is, when the program has not been implemented as planned (also called program or treatment fidelity; Bellg, et al., 2005; Dusenbury, Brannigan, Falco, & Hansen, 2003).

Low fidelity to program content is linked with poor outcomes (Bellg, et al., 2005; Mihalic & Irwin, 2003; Pentz, 2003). In fact, poor implementation can sometimes be the cause of the failure of a program (Han & Weiss, 2005). Small effect sizes or findings inconsistent with well-reasoned hypotheses may not be related to the efficacy of the program as it was designed, but rather be related to failure to implement the program as intended. Thus, in the absence of evaluation of the fidelity of implementation, it is difficult to disentangle whether a weak effect size or an unexpected outcome was due to poor program conceptualization and design or to poor program implementation. In fact, without process evaluation, evaluators may never know if the intervention was implemented at all.

Furthermore, evaluation of program implementation gives insight into the process of the intervention (Domitrovich & Greenberg, 2000). Such insight is helpful in optimizing program quality and efficiency. If the most faithful implementation of the program produces the most favorable outcomes, program administrators and evaluators will have greater confidence that the intervention is responsible for the change (Domitrovich & Greenberg, 2000). According to Dusenbury, Brannigan, Falco, & Hansen (2003), fidelity is measured by adherence to program, dose (amount of program delivered), quality of program delivery, participant responsiveness, and program differentiation (whether critical features that distinguish the program are present). Training those who deliver the program is a critical issue, although one that is not well documented (Bellg, et al., 2005; Dusenbury, Brannigan, Falco, & Hansen, 2003). Therefore, a critical part of program evaluation should include not only outcome data, but also process data that address these issues of program implementation.

Process Evaluation Data

Educational and social interventions may be thought of as processes because they involve acts that are intended to create outcomes that are of value to constituents (e.g., Pyzdek, 2003). Process evaluation involves measurement, analysis, and interpretation of data that reflect how a program or intervention was delivered—that is, the procedures or process used during implementation. In the case of an educational intervention, for example, process evaluation might address the efficacy of different teaching methods, laboratory or homework assignments, and written resources. Program fidelity is reflected in the extent to which these components were executed in a manner that is consistent with the intentions of the program leaders. Examples of evaluation questions that address program fidelity include the following: “Did the teachers, in fact, assign the designated homework to the students?” and, “Did the teachers cover all facets of the lesson as designed in the lesson plan?”

Such process data are becoming increasingly popular tools to assess program implementation and fidelity. The literature, however, has mixed reports of evidence-

based programs being implemented with fidelity. One review of program implementation reported that some evidence-based violence prevention programs were implemented as planned between 57% and 74% of the time (Elliott & Mihalic, 2004). A popular drug prevention program (Life Skills Training, e.g., Botvin, Mihalic, & Grotmeter, 1998) was implemented with fidelity approximately 83% of the time, although Botvin, Mihalic, & Grotmeter (1998) have reported lower levels of fidelity, at around 49% (Elliott & Mihalic, 2004). Elliot and Mihalic (2004) reported other drug (including smoking) prevention programs, as well as other “healthy lifestyles” and bullying curricula, have been implemented with fidelity between 24% and 55% of the time.

Linking process data with outcome data, however, is relatively rare, perhaps in part because of the methodological challenges associated with directly linking specific content and activities with specific outcomes. Nevertheless, it seems a critical issue to understand how a conceptually sound program may be linked to poor outcomes. Process data are also important in determining which elements of a program are critical or more important than others, that is, researchers and program providers will know how faithfully the intervention must be implemented in order to produce the desired outcomes. Likewise, process data can also reveal which elements of a program may be least effective in producing the desired outcomes. This is a particularly important point, as the amount of time devoted to “human service” programs in an in-school or after-school environment is often limited. Linking process data with outcomes may also shed light on how to increase fidelity in the future (Dusenbury, Brannigan, Falco, & Hansen, 2003; Elliot & Mihalic, 2004; Gingiss, Roberts-Grey, & Boerm, 2006). Finally, understanding process also provides valuable information regarding the sustainability and feasibility of the intervention once researchers and government funding are no longer available.

Background on HealthWise and the Motivation Lesson

The intervention of interest to this paper is HealthWise South Africa, which aims to prevent risky behavior among adolescents (in particular, risky sexual behavior and substance misuse). Funded by the National Institute on Drug Abuse, it is being evaluated through a three-cohort, multi-wave randomized control trial. HealthWise combines elements of Life Skills Training (Botvin, Mihalic, & Grotmeter, 1998; Botvin & Kantor, 2000), TimeWise: Taking Charge of Leisure Time (Caldwell, 2004), and lessons drawn from effective sexual risk prevention curricula to develop a comprehensive risk reduction curriculum based on a positive youth development approach. Thus, most of the 18 lessons are contextualized, if not focused specifically, on leisure. Extensive process and outcome data have been collected to date, although data collection is still on-going.

Curriculum Development and Description

HealthWise is being implemented in four schools as part of the Life Orientation Curriculum, which is a mandated subject in South African schools. Each HealthWise

lesson objective has been mapped onto the specific educational standards outlined in the Life Orientation Curriculum.

Table 1. HealthWise South Africa Curriculum Lessons

Grade 8
Lesson 1: Self-awareness
Lesson 2: Managing anxiety
Lesson 3: Managing anger
<i>Lesson 4: Exploring free time</i>
<i>Lesson 5: Free time in my community</i>
<i>Lesson 6: Beating boredom and developing interests</i>
<i>Lesson 7: Overcoming roadblocks</i>
<i>Lesson 8: Decision making</i>
Lesson 9: Managing risk
Lesson 10: Avoiding risky sexual behavior
Lesson 11: Myths & realities of drug use
Lesson 12: Avoiding and reducing risk
Grade 9
Lesson 1: Review
<i>Lesson 2: Leisure motivation</i>
<i>Lesson 3: Community connections</i>
<i>Lesson 4: Planning and managing leisure</i>
Lesson 5: Relationships and sexual behavior
Lesson 6: Conflict resolution

^a Italics indicate leisure lessons

As shown in Table 1, the HealthWise curriculum is given in 8th and 9th grades; 12 lessons are delivered in 8th grade and six in 9th grade. Each lesson takes approximately three class sessions. Although there are some specific leisure lessons (shown in italics in the table), other material is often based on leisure content. For example, in lesson three (learning to deal with anger), learners are asked to imagine themselves on a soccer field where they get into a conflict. Or, in lesson 12 (learning to avoid risk), learners role play being in leisure-based situations that can turn risky. Thus, HealthWise both directly and indirectly addresses how leisure can lead to both healthy and unhealthy outcomes. For more detail on the development of the curriculum, see Caldwell et al. (2004).

Theoretical Foundation for the Motivation Lesson. One of the key theoretical bases for HealthWise is self-determination theory (SDT; Ryan & Deci, 2000a, 2000b). The role of motivation in adolescent development has been a growing topic of research interest, particularly since different forms of motivation are associated with positive outcomes and others are associated with negative outcomes. In particular, motivation that is more self-determined is linked with positive outcomes such as psychological well-being (Larson, 2000; Ryan & Deci, 2000a, 2000b). On the other hand, amotivation and more extrinsic forms of motivation are linked with unhealthy outcomes such as boredom (Baker, 2004; Standage, Duda, & Ntoumanis, 2005; Walker, Greene,

& Mansell, 2006). Thus, motivation can be considered a risk factor or a protective factor.

SDT is often used to understand why humans do what they do, as well as the concomitant outcomes of behavior (Ryan & Deci, 2000a, 2000b). It is particularly important to understand reasons that compel adolescents to act because youth at this stage are learning to become increasingly independent in thought and action from their parents, in addition to being faced with often enormous amounts of peer pressure to enact both positive and negative behaviors. At the same time, adolescents are learning how to assert their own personal views and identities in light of the influence from parents, family, and peers (Steinberg, 1996). From a risk and prevention perspective, research suggests that at least for U.S. samples, type of leisure motivation is associated with either decreased or increased risk for substance use and other risky behavior (e.g., Caldwell & Darling, 1999; Caldwell, Darling, Payne, & Dowdy, 1999; Iso-Ahola & Crowley, 1991).

SDT posits that humans are motivated by rewards internal to the self (e.g., “I do it because it pleases me”) or external to the self (e.g., “I have to do it due to coercion or opportunity for a significant reward that is not intrinsic to the task”); Ryan & Deci, 2000a, 2000b). SDT suggests that motivation can be understood on a continuum, with *fully intrinsically motivated* behavior anchoring one end. In the middle of the continuum are two different motivational orientations: *identified* regulation is characterized by pursuit of a goal, which is personally important but not fully owned, and *introjected* regulation, which is characterized by behavior driven by perceptions of what others might think. Anchoring the other end of the continuum is *external* regulation, which is behavior compelled due to external influence or done to gain a reward that has nothing to do with the intrinsic value of the behavior. *Amotivation*, which falls outside of the continuum, is characterized by no motivation at all: that is, an individual cannot provide a reason for his or her behavior. There is a sense of alienation and lack of control, and actions are performed without purpose or volition.

Motivation can occur at the global (personality), contextual (domains such as work, education, or leisure), or situational (specific activity or state) levels (Vallerand, 1997). Vallerand has proffered a hierarchical model to depict these levels and suggested that there are recursive influences of motivation among the levels. Thus, motivation exists both within an individual and in relation to his or her interactions with the environment (Ryan & Deci, 2000a). We choose to conceptualize and measure motivation at the contextual level because the other variables in the larger study were also measured at the contextual level (e.g., boredom in leisure, decision making and planning in leisure). HealthWise also was conceptualized at the contextual level, as the intention of the intervention was to assist adolescents in being responsible for managing the whole of their leisure time. Thus, conceptual congruence between measurement and intervention focus was achieved across the study.

With regard to a risk and protective framework, leisure motivation is conceptualized to mediate the effects of the intervention on mitigating risky behavior. Leisure time is a major context to support intrinsically motivated behavior for adolescents (Caldwell, 2005; Larson, 2000) because it maximizes the opportunity to engage in activity out of interest and personal choice. Theoretically, those with high levels of amotivation would be more likely to engage in risky or negative behaviors since they would be

more likely to be disengaged and unattached to prosocial norms, or would be more influenced by peers. Those engaged in an activity for a purpose (identified motivation) or for its intrinsic value (e.g., it is enjoyable) would be less likely to engage in risky behavior, as they would likely be engaged in personally meaningful and fulfilling leisure (Larson, 2000). In addition, they are more likely to be attached and bonded to prosocial institutions and norms (e.g., Eccles & Barber, 1999). The role of introjected motivation, which implicates peer influence, is more difficult to predict because in a leisure context, peers can exert both positive and negative influences on behavior or experience (Caldwell & Baldwin, 2005). Thus, one of the goals of HealthWise is to increase internalized forms of motivation (intrinsic and identified) among adolescents and decrease amotivation in their leisure.

Most of the research conducted on adolescent motivation has been correlational in nature and has begged the question of whether an intervention can change adolescents' typical motivational disposition from a more amotivated or external structure to a more internal reward structure. Intrinsic motivation, however, may be difficult to influence (e.g., Digelidis, Papaioannou, Laparidis, & Christodoulidis, 2003). In a previous study on rural U.S. youth, for example, relative to a comparison group, intervention participants reported lower levels of amotivation and greater identified and introjected self-regulation, with small to moderate effect sizes (Caldwell, Baldwin, Walls, & Smith, 2004). On the other hand, there were no statistically significant effects on intrinsic and extrinsic motivation.

HealthWise Lesson Two: Leisure Motivation. In HealthWise, leisure motivation is addressed primarily in lesson two in 9th grade. In this lesson, learners are introduced to reasons associated with amotivated, externally, and internally motivated styles of leisure behavior. The lesson focuses on identifying activities that learners engage in because they have a real interest or because the activity may serve a future purpose, such as learning to play an instrument to get into the school band (identified motivation). Learners are taught that more benefits accrue if they do things in their leisure time that are inherently interesting. In contrast, activities that are done because there is nothing else to do (amotivation), because there is coercion or reward involved (external motivation), or because there is a need to fit in or to be popular with friends (introjected motivation) are also discussed.

Educator Training

Educator training sessions were held three weeks prior to the beginning of each school year to train educators in the upcoming HealthWise lessons. Four of the educators had participated in a pilot study and had formerly attended a training session. In 2004, 12 of the 18 invited educators took part in a two-day training workshop. In 2005, 10 of the 25 educators (all from School A) took part in the formal training workshop; of the 15 who didn't attend, eight had received training in previous years. The Cape Town research team met for an informal training session with the seven educators who did not attend the formal training workshop. During the training workshop, trainers described the curriculum in detail and educators participated in role-plays and practiced teaching various lessons. Trainers also instructed educators on how to complete the process evaluation forms and were asked to complete them after the delivery of each lesson.

Method

Sample

The sample for this study is comprised of South African youth who are living in a low-income township outside of Cape Town, South Africa. Overall, we are collecting data on three sequential cohorts of learners from nine schools. Four schools serve as intervention schools ($n = 902$, 41.1%) and five schools serve as control schools ($n = 1291$, 58.9%). Data collection began in February 2004, and will end in October 2008.

For each cohort, baseline data are collected on learners as they begin 8th grade. Currently four waves of data have been collected on the first cohort, which is the focus of this paper. The mean age of the sample at wave 3 was 15.0 years ($SD = .86$) and 51% of students were female. Most participants (85.8%) identified themselves as Colored (a combination of Asian, European, and African descent) with an additional 9.5% of students identifying as Black and 4.0% as White. Most students were Christian (23.2% Catholic, 44.8% Other Christian), with 27.9% reporting their religion as Islam. Language spoken at home was mixed: 52.7% spoke English, 45.6% spoke Afrikaans, and 5.1% spoke Xhosa. Given the mix of languages, the program materials are available in both English and Afrikaans.

Data Collection

Two types of data were collected in this study. Outcome data were collected through self-report questionnaires. Process data were collected through use of a checklist completed by program staff.

Outcome Evaluation Procedures. Learner questionnaires were self-administered through the use of palm pilots at the beginning (wave 1, February 2004) and end (wave 2, October 2004) of 8th grade and the beginning (wave 3, February 2005) and end (wave 4, October 2005) of 9th grade. HealthWise was delivered between waves 1 and 2 and between waves 3 and 4. The lesson on motivation was taught during 9th grade. Therefore, of particular interest are the potential changes in motivation between waves 3 and 4.

Implementation Evaluation Procedures. In order to measure program fidelity, a one-page form was developed to address the elements that reflect program implementation (Dusenbury, Brannigan, Falco, & Hansen, 2003): adherence, dose, quality of delivery, participant responsiveness, and program differentiation. Educators were asked to answer each set of questions after each lesson taught.

Measures

Of interest to this paper were measures of how the motivation lesson was implemented and learner reaction to it. We also measured motivation using a general scale with five sub-scales.

Motivation Outcome Measure. The Free Time Motivation Scale for Adolescents (FTMS-A) was developed to reflect the SDT framework in a free time context among adolescents approximately aged 12 to 16 years (Baldwin & Caldwell, 2003). It is a measure of motivation at the contextual (i.e., leisure) level. Acceptable reliability and evidence of validity were established among a sample of rural adolescent (Baldwin & Caldwell). The five sub-scales of the FTMS-A were coded on a five point Likert-type scale anchored by 0 (low level of the variable) and 4 (high level of the variable).

These sub-scales included amotivation (3 items, $\alpha = .72, .79, .80, .83$, waves 1 through 4, respectively; e.g., “I don’t know why I do my free time activities, nothing much interests me”), external motivation (3 items, $\alpha = .76, .75, .82, .81$; e.g., “I do what I do in my free time because my parents expect me to”), introjected motivation (3 items, $\alpha = .67, .69, .73, .71$; e.g., “I do what I do in my free time because I want to impress my friends”), identified motivation (6 items, $\alpha = .78, .79, .81, .79$; e.g., “I do what I do in my free time because it is important to me”), and internal motivation (2 items, $\alpha = .58, .65, .64, .67$; e.g., “I do what I do in my free time because I want to have fun”). Reliability generally increases as learners mature, which may reflect their familiarity with the measures or may reflect a developmental process as adolescents increase in their capacity to reflect on their motivations and internal states as they mature.

Implementation and Fidelity Data. The implementation evaluation form was made up of the following items assessing the different elements that measure program implementation:

1. What was the learners’ general response to the content? (participant responsiveness; response format ranged from 1 = very negative to 5 = very positive)
2. What was the learners’ general response to the teaching methods? (participant responsiveness; response format ranged from 1 = very negative to 5 = very positive)
3. What was the learners’ general response to the worksheets? (participant responsiveness; response format ranged from 1 = very negative to 5 = very positive)
4. How many learners were involved? (dose; response format ranged from 1 = none of the learners to 5 = almost all of the learners)
5. How much content was covered? (dose, adherence and quality of delivery; response format was 1 = 0% - nothing, 2 = 25% - not much, 3 = 50% - half of it, 4 = 75% - most of it, and 5 = 100% - all of it).
6. What comments do you have about the lesson and how the students responded? (potentially any of the factors).

Program differentiation was assessed by whether or not the educator actually taught the lesson at all. For the purposes of this paper, only the one-page evaluation form for the motivation lesson is used in the analysis.

Analysis

Process Data and Outcome Data. Two analyses were of particular interest. First, we were interested in comparison of the process measures from the educator-completed check sheets across the four schools. Because educators in School A received the most extensive training, we anticipated scores on the process measures from that school to be significantly higher than scores from the other three schools.

We were also interested in contrasting outcome means (motivation types) of wave 1 with wave 4, and wave 3 with wave 4. The wave 1 to 4 contrast was important because learners are exposed to HealthWise leisure lessons across both years, and it is reasonable to expect that HealthWise may have indirectly influenced learners’ motivation. The HealthWise motivation lesson was given between waves 3 and 4. As

such, we expected that involvement in the intervention processes that directly targeted motivational outcomes would produce the strongest effect on motivation. In terms of program implementation and fidelity, the contrast of wave 1 with wave 4 is largely an evaluation of the “dose” effect, while the contrast of wave 3 with wave 4 is an evaluation of program differentiation (Dusenbury, Brannigan, Falco, & Hansen, 2003).

Our analysis was also guided by observations and interactions with educators by the research team in Cape Town. It appeared that educators in one of the treatment schools (School A) were more motivated to complete the process forms, more engaged with the curriculum, and had more principal support. In addition, as previously mentioned, they had the highest level of training. All of these issues relate directly to implementation fidelity and would indicate that the curriculum was being better implemented in School A than in the other treatment schools.

Based on this a-priori belief, we analyzed the outcome data two ways in order to better address our hypotheses that learners in the treatment schools would increase in intrinsic and identified motivation and decrease in amotivation and external motivation. In addition to comparing all treatment schools (A, B, C, and D) to the control group, School A only is compared to the control group. In this way, we attempted to uncover whether better implementation as suggested by the process data may produce better outcomes.

Outcome data were analyzed using a repeated measures general linear model (GLM), where time is the within subjects factor. We also calculated effect sizes using Cohen’s (1988) *d*. This analysis strategy allowed us to examine change in motivation across time, while controlling for previous levels of motivation.

Linking Process and Outcome Data. One of the complications encountered while analyzing the process data resulted from the non-uniform manner in which the educators completed the process evaluation forms. In School A, five educators, each teaching one classroom, completed the motivation lesson process evaluation form (i.e., five different classes received HealthWise in School A and five evaluation forms were submitted). In School B, two educators each taught three classes (for a total of six HealthWise classes). But these educators averaged their scores and submitted only one form each (i.e., per teacher) per lesson, thus only two evaluation forms were submitted. In School C, one educator taught four HealthWise classes and completed one form for each lesson. In School D, one educator taught multiple HealthWise classes but completed only one form. For the motivation lesson, each process evaluation item was answered with the highest rating. Thus, School D was omitted from the implementation and fidelity analysis.

The unevenness in the way in which educators completed the process evaluation forms posed a challenge in trying to map the process evaluation data onto the outcome data. The research team decided to do the following to have a basic point of comparison across the three schools: Within each school, all teachers’ evaluations were averaged, such that there is one score on each variable, per school. Thus, there were five educator responses in School A, and six for Schools B and C. Although School D was omitted from analysis, we did include School D’s outcome data, as the main goal was to isolate the effect of School A versus the other schools. Although perhaps not ideal, it seemed the best way to handle the situation and not lose potentially important process data. Due to low numbers of participating teachers, process data were analyzed by visual inspection as well as by their effect sizes.

Results

As can be seen in Table 2, our initial hunches about differences in program implementation were supported based on a visual inspection of the data as well as the effect sizes, although *t*-tests did not reveal any statistically significant differences. The responses by learners to the teaching methods and lesson material differed to some degree, and the .37 effect size (*d*) for both comparisons indicates a somewhat moderate difference between School A and Schools B and C. A large effect size was found for the dose and quality of implementation (*d* = 1 for both). Compared with educators in Schools B and C, educators in School A reported covering more content of the motivation lesson (mean = 4.80 compared with 3.83, respectively) and indicated a higher number of learners who were exposed to the lesson (mean = 4.60 compared with 3.67, respectively). It is also interesting to note that of the five comments from educators in School A, three mentioned that their learners lacked motivation, implying that the educators may have had a heightened awareness of the importance of motivation.

Table 3 displays the outcome data associated with the lesson on motivation. This table is organized in the following manner. Each type of motivation is addressed separately. Data are reported from Waves 1, 3, and 4. Wave 1 was the baseline, pre-test data. The lesson specific to motivation was given between Waves 3 and 4, although other lessons in the curriculum indirectly addressed leisure motivation. Wave 2 data did not add any new

Table 2. Process Evaluation Data

Variable	School	N	Mean (s.d.)	Cohen's <i>d</i>
Learners' general response ¹	School A	5	3.80 (0.45)	.37
	Schools B&C	6	3.50 (1.05)	
Response to teaching methods ¹	School A	5	3.80 (0.45)	.37
	Schools B&C	6	3.50 (1.05)	
Response to worksheets ¹	School A	5	3.80 (0.45)	.03
	Schools B&C	6	3.83 (0.75)	
Number of learners involved ²	School A	5	4.60 (0.55)	1.0
	Schools B&C	6	3.67 (1.20)	
Content covered ³	School A	5	4.80 (0.45)	1.0
	Schools B&C	6	3.83 (1.17)	
Open ended comments	School A	5	My students lack motivation (N=3). Comprehension was not made (N=1). The lesson was thought provoking (N=1).	
	Schools B&C	6	Found the activity boring (N=1). Liked the activity (N=1).	

¹Response format: 1 = very negative to 5 = very positive

²Response format: 1 = none of the learners to 5 = all of the learners

³Response format: 1 = 0%, 2 = 25%, 3 = 50%, 4 = 75%, 5 = 100%

information to the table and just added another level of complexity; thus, it was omitted. Means and standard errors are reported for each wave of data collection.

Given the complexity of the analyses, the last column of Table 3 summarizes the findings by reporting the F - and p -values for each significant GLM analysis and associated effect sizes (d). For amotivation, we found that there was a treatment effect. Compared to the control schools, learners' levels of amotivation in all treatment schools decreased from wave 3 to wave 4 ($F = 8.70, p = .003$) as hypothesized, although

Table 3. Results of repeated measures GLM across three waves by type of motivation: Comparisons by all treatment schools, treatment school A, and treatment schools BCD

AMOTIVATION				
Wave	Treatment Mean (s.e.)		Control Mean (s.e.)	Significance
	A, B, C, D	A		ABCD vs. Control W3 vs. W4, $F = 8.70, p = .003, d = .16$ A vs. Control W1 vs. W4, $F = 5.26, p = .022, d = 1.43$ W3 vs. W4, $F = 10.12, p = .003, d = .20$
1	1.69 (.042)	1.72 (.075)	1.56 (.034)	
3	1.82 (.045)	1.77 (.076)	1.55 (.035)	
4	1.71 (.045)	1.55 (.079)	1.62 (.036)	
EXTRINSIC MOTIVATION				
Wave	Treatment Mean (s.e.)		Control Mean (s.e.)	Significance
	A, B, C, D	A		
1	2.01 (.047)	2.13 (.083)	1.99 (.037)	
3	1.81 (.048)	1.81 (.083)	1.62 (.038)	
4	1.67 (.046)	1.66 (.080)	1.51 (.036)	
INTROJECTED MOTIVATION				
Wave	Treatment Mean (s.e.)		Control Mean (s.e.)	Significance
	A, B, C, D	A		A vs. Control W1 vs. W4, $F = 5.50, p = .019, d = .15$
1	1.74 (.041)	1.81 (.071)	1.60 (.032)	
3	1.66 (.044)	1.68 (.076)	1.52 (.035)	
4	1.55 (.041)	1.49 (.073)	1.48 (.033)	
IDENTIFIED MOTIVATION				
Wave	Treatment Mean (s.e.)		Control Mean (s.e.)	Significance
	A, B, C, D	A		A vs. Control W1 vs. W4, $F = 4.10, p = .043, d = .13$
1	2.78 (.025)	2.81 (.057)	2.79 (.025)	
3	2.66 (.035)	2.74 (.059)	2.68 (.027)	
4	2.64 (.035)	2.79 (.060)	2.62 (.027)	
INTRINSIC MOTIVATION				
Wave	Treatment Mean (s.e.)		Control Mean (s.e.)	Significance
	A, B, C, D	A		A vs. Control W1 vs. W4, $F = 5.70, p = .017, d = .15$
1	2.88 (.039)	2.84 (.037)	3.01 (.030)	
3	2.88 (.039)	2.91 (.039)	2.98 (.030)	
4	2.91 (.039)	3.00 (.038)	2.97 (.030)	

Note: Responses coded on a scale from 0 to 4, where 0 indicates a lower level of the variable.

the effect size was small ($d = .16$). Comparing School A only to the control schools learners, there is a large wave 1 to wave 4 difference ($F = 5.26, p = .022, d = 1.43$) and a wave 3 to wave 4 difference ($F = 10.12, p = .002$), although the effect size is small ($d = .20$).

Across the other types of motivation, there were no other “all treatment schools” (i.e., A, B, C, and D) versus control school differences. There were, however, three School A versus control schools differences for introjected, identified, and intrinsic motivation. Compared to the control schools, learners in School A reported lower levels of introjected motivation between waves 1 and 4. Their levels of identified motivation remained relatively stable as compared to learners in the control schools, whose levels of identified motivation declined between waves 1 and 4. Finally, their levels of intrinsic motivation slightly increased compared to the control learners’ levels, which slightly decreased between waves 1 and 4. All of the changes are in the desired direction although the effect sizes were small (d ranged from .13 to .15).

Discussion

This study sought to examine the role of implementation fidelity in understanding leisure motivation outcomes. Level of training, amount of content covered, quality of implementation, response of participants, and dose were considered.

Based on the first cohort of learners exposed to HealthWise, it does appear that there may be an overall program effect of the curriculum on amotivation regardless of fidelity of implementation. Compared to the control schools, all treatment school learners reported lower levels of amotivation in wave 4 compared to wave 3, and wave 4 compared to wave 1 as hypothesized. Using process evaluation data to monitor implementation fidelity, however, we also conclude that the school with better trained teachers who also reported higher levels of program fidelity had a stronger influence on decreasing amotivation between waves 1 and 4 than the other schools.

The effect of HealthWise on changing levels of the other types of motivation does not seem as robust, however. Although there are no other all treatment schools versus control schools differences, there are School A versus control schools differences in the hypothesized direction for identified and intrinsic motivation between waves 1 and 4. There was also a statistically significant difference for introjected motivation (School A means significantly decreased between waves 1 and 4 compared to control schools).

Taken together, these statistically significant differences and what we observed from process data (including differential training levels) suggest that program fidelity may be linked to outcomes. If we had not taken into account training and program implementation and fidelity issues, we would not have looked for or seen changes in levels of introjected, identified and intrinsic motivation between Waves 1 and 4. Although most of the effect sizes were small, differences did exist that would have gone undetected. Small effect sizes are more common in social science research and are influenced by a great many factors, such as measurement error, study design, and potency of the intervention, and thus should not be dismissed as trivial (Cohen, 1988; McCartney & Rosenthal, 2000). We did, however, find a large effect size for amotivation when comparing School A with the control schools from wave 1 to wave

4, again suggesting that had we not taken into account process data, we would not have uncovered this particularly strong finding.

Although there is no research that has focused on which type of motivation, or combinations of motivation, are related to greater levels of risk or protection, amotivation is associated with negative outcomes (Sheldon, Ryan, Deci, & Kasser, 2004). This is particularly important because middle adolescents' levels of amotivation naturally increase across this developmental period (e.g., Sharp, Caldwell, Graham, & Ridenour, 2006). There is a concomitant decline in more intrinsic forms of motivation (e.g., Sharp, Caldwell, Graham, & Ridenour, 2006). Thus, an intervention that not only can stave off the developmental trajectory of increased amotivation but can also decrease levels of it holds promise in terms of preventing risky outcomes. Because HealthWise seemed to help students become less amotivated regardless of level of implementation and fidelity, the curriculum seems robust with regard to amotivation, although much more research needs to be conducted to conclude this with any certainty.

We were surprised that there were no wave 3 to wave 4 differences for the motivation types other than amotivation. As future waves of data come in for cohort one, as well as the other two cohorts, we will continue to monitor the effects of the intervention on motivation. For now, we can conclude that our attempts to influence leisure motivation in general hold some promise, but a much stronger dose or potent intervention might be needed.

Limitations

Our analysis is limited by the quality of the process data we were able to collect for this first cohort. Nevertheless, we felt that we could learn something from these analyses. We have stressed to the educators the need to complete the process evaluation forms for each class taught. Another way to collect this data in a perhaps less subjective way would be through observation of the lessons. Unfortunately, the educators in this study felt uncomfortable being observed, which prevented our ability to collect this type of data as we had hoped. We also have concluded that it would be helpful to have more specific information on each lesson in terms of the specific activities that were successfully completed. Thus, we have modified the form to be more user friendly and to address specific activities within each lesson.

Another concern is that even if we had high quality process data, the methodology to link process data with outcomes is lacking. Advanced, multi-level methods are needed to address issues of educators and learners being nested within schools and classrooms. Because this is an area just receiving attention in the evaluation literature, we anticipate increased attention to these methodological issues.

Conclusion

In this paper, we addressed only one outcome of a complex, multi-outcome, multi-cohort study with potentially seven waves of data collected on the first cohort alone. The process of linking program implementation and fidelity data with these multiple outcomes within cohort, let alone across cohorts, is a challenging prospect. We have highlighted and illustrated some of the issues related to linking process and outcome

data with this first cohort of HealthWise learners. As methods to improve linking process data with outcomes become more sophisticated, we hope to be able to shed more light on how implementation can affect outcomes so that program implementation can be improved and program effects can be better interpreted.

References

- Baker, S. R. (2004). Intrinsic, extrinsic, and amotivational orientations: Their role in university adjustment, stress, well-being, and subsequent academic performance. *Current Psychology: Developmental, Learning, Personality, Social*, 23, 189-202.
- Baldwin, C. K., & Caldwell, L. L. (2003). Development of the free time motivation scale for adolescents. *Journal of Leisure Research*, 35, 129-151.
- Bellg, A. J., Borelli, B., Resnick, B., Hecht, J., Minicucci, D. S., Ory, M., Ogedegbe, G., Orwig, D., Ernst, D., Dzajkowski, S. (2004). Enhancing treatment fidelity in health behavior change studies: Best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychology*, 23, 443-451.
- Botvin, G., Mihalic, S. F., & Grotpeter, J. K. (1998). Life skills training. In D. S. Elliott (Series Ed.), *Blueprints for Violence Prevention*, Book 5, Center for Study and Prevention of Violence, Institute of Behavioral Science, University of Colorado, Boulder, CO.
- Botvin, G. J., & Kantor, L. W. (2000). Preventing alcohol and tobacco use through life skills training. *Alcohol Research and Health*, 24, 250-257.
- Caldwell, L. L. (2005). Recreation and youth development. In P. A. Witt & L. L. Caldwell (Eds.), *Recreation and Youth Development* (pp. 169-192). State College, PA: Venture
- Caldwell, L. L., & Baldwin, C. K. (2005). What constrains adolescents' leisure? A developmental approach. In E. Jackson (Ed.), *Constraints to Leisure* (pp. 75-88). State College, PA: Venture.
- Caldwell, L. L., Baldwin, C. K., Walls, T. & Smith, E. A. (2004). Preliminary effects of a leisure education program to promote healthy use of free time among middle school adolescents. *Journal of Leisure Research*, 36, 310-335.
- Caldwell, L. L., & Darling, N. (1999). Leisure context, parental control, and resistance to peer pressure as predictors of adolescent partying and substance use: An ecological perspective. *Journal of Leisure Research*, 31, 57-77.
- Caldwell, L. L. (2004). *TimeWise: Taking charge of leisure time curriculum for middle school students*. Scotts Valley, CA: ETR Associates.
- Caldwell, L. L., Darling, N., Payne, L., & Dowdy, B. (1999). "Why are you bored?" An examination of psychological and social control causes of boredom among adolescents. *Journal of Leisure Research*, 31, 103-121.
- Caldwell, L. L., Smith, E. A., Wegner, L., Vergnani, T., Mporfu, E., Flisher, A., & Matthews, C. (2004). HealthWise South Africa: Developing a life skills curriculum for young adults. *World Leisure Journal*, 46(3), 4-17.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.) Hillsdale, NJ: Lawrence Earlbaum Associates.

- Digelidis, N., Papaioannou, A., Lapidis, K., & Christodoulidis, T. (2003). A one-year intervention in 7th grade physical education classes aiming to change motivational climate and attitudes toward exercise. *Psychology of Sport and Exercise, 4*, 195-210.
- Domitrovich, C. E., & Greenberg, M. T. (2000). The study of implementation: Current findings from effective programs that prevent mental disorders in school-aged children. *Journal of Educational and Psychological Consultation, 11*, 193-221.
- Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. *Health Education Research, 18*, 237-256.
- Eccles, J. S., & Barber, B. L. (1999). Student council, volunteering, basketball, or marching band: What kind of extracurricular involvement matters? *Journal of Adolescent Research, 14*, 10-43.
- Elliott, D. S., & Mihalic, S. (2004). Issues in disseminating and replicating effective prevention programs. *Prevention Science, 5*, 47-53.
- Gingiss, P. M., Roberts-Gray, C., & Boerm, M. (2006). *Bridge-It*: A system for predicting implementation fidelity for school-based tobacco prevention programs. *Prevention Science, 7*, 197-207.
- Guskey, T. R. (2000). *Evaluating Professional Development*. Thousand Oaks, CA: Corwin Press, Inc.
- Han, S. S., & Weiss, B. (2005). Sustainability of teacher implementation of school-based mental health programs. *Journal of Abnormal Child Psychology, 33*, 665-679.
- Iso-Ahola, S. E., & Crowley, E. D. (1991). Adolescent substance abuse and leisure boredom. *Journal of Leisure Research, 23*(3), 260-271.
- Larson, R. (2000). Toward a psychology of positive youth development. *American Psychologist, 55*, 170-183.
- McCartney, K. & Rosenthal, R. (2000). Effect size, practical importance, and social policy for children. *Child Development, 71*, 173-180.
- Mihalic, S., & Irwin, K. (2003). Blueprints for violence prevention: From research to real-world settings—Factors influencing the successful replication of model programs. *Youth Violence and Juvenile Justice, 1*, 307-329.
- Pentz, M. (2003). Evidence-based prevention: Characteristics, impact, and future direction. *Journal of Psychoactive Drugs, 35*, 143-152.
- Pyzdek, T. (2003). *The Six Sigma handbook*. New York: McGraw-Hill.
- Ryan, R. M., & Deci, E. L. (2000a). Intrinsic and extrinsic motivation: Classic definitions and new directions. *Contemporary Educational Psychology, 25*, 54-67.
- Ryan, R. M., & Deci, E. L. (2000b). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*, 68-78.
- Sharp, E. H., Caldwell, L. L., Graham, J. W., & Ridenour, T. A. (2006). Individual motivation and parental influence on adolescents' experiences of interest in free time: A longitudinal examination. *Journal of Youth and Adolescence, 35*, 340-353.
- Sheldon, K. M., Ryan, R. M., Deci, E. L., & Kasser, T. (2004). The independent effects of goal contents and motives on well-being: It's both what you pursue and why you pursue it. *Personality and Social Psychology Bulletin, 30*, 475-486.

- Standage, M., Duda, J. L., Ntoumanis, N. (2005). A test of self-determination theory in school physical education. *British Journal of Educational Psychology*, 75, 411-433.
- Steinberg, L. (1996). *Adolescence* (4th ed.). New York: McGraw-Hill.
- Vallerand, R. J. (1997). Toward a hierarchical model of intrinsic and extrinsic motivation. In M. P. Zanna (Ed.), *Advances in experimental social psychology*, 29, 271-360.
- Walker, C. O., Greene, B. A., & Mansell, R. A. (2006). Identification with academics, intrinsic/extrinsic motivation, and self-efficacy as predictors of cognitive. *Learning and Individual Differences*, 16, 1-12.

Copyright of Journal of Park & Recreation Administration is the property of Sagamore Publishing, LLC and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.