Using Students' Weekly Diaries to Evaluate Positive Youth Development Programs:

Are Findings Based on Multiple Studies Consistent?

Daniel T.L. Shek

Department of Applied Social Sciences
The Hong Kong Polytechnic University

Running Head: Students' weekly diaries

Abstract

Asking clients to document their perceived quality of life during and after intervention is a popular approach employed by helping professionals to evaluate intervention programs. In the Project P.A.T.H.S., students participating in the Experimental Implementation Phase and Full Implementation Phase were invited to write reflective journals in the form of weekly diaries to reveal their perceptions and feelings regarding the program and the perceived benefits of the program. Based on multiple studies, results showed that the respondents generally (a) had positive views on the program, (b) had positive views on the instructors, and (c) perceived that they had acquired competencies at the societal, familial, interpersonal and personal levels and their quality of life was promoted after joining the program. Acknowledging the limitations of diaries, the present qualitative findings provide support for the effectiveness of the Tier 1 Program of the Project P.A.T.H.S. in Hong Kong.

Using Students' Weekly Diaries to Evaluate Positive Youth Development Programs:

Are Findings Based on Multiple Studies Consistent?

The main purpose of this paper is to report evaluation findings of the Project P.A.T.H.S. which is a positive youth development program in Hong Kong. To promote holistic development among adolescents in Hong Kong, The Hong Kong Jockey Club Charities Trust has approved HK$400 million to launch a project entitled "P.A.T.H.S. to Adulthood: A Jockey Club Youth Enhancement Scheme". The word "P.A.T.H.S." denotes **P**ositive **A**dolescent **T**raining through **H**olistic **S**ocial Programmes. There are two tiers of programs (Tier 1 and Tier 2 Programs) in this project. The Tier 1 Program is a universal positive youth development program in which students in Secondary 1 to Secondary 3 will participate, normally with 20 hours of training in the school year at each grade. Because research findings suggest that roughly one-fifth of adolescents would need help of a deeper nature, the Tier 2 Program will generally be provided for at least one-fifth of the students who have greater psychosocial needs at each grade (i.e., selective program). In this paper, three qualitative studies based on students' diaries were conducted to understand the perceived attributes of the Tier 1 Program and perceived benefits of the Tier 1 Program on the quality of life (such as psychosocial competencies) of the program participants.

Several characteristics are intrinsic to the curriculum design of the Tier 1 Program. First, although the number of hours for each grade of the junior secondary schools is 20, schools with special needs may choose to focus on the core units only (i.e., 10 hours for the Tier 1 Program), so that they can allocate manpower and resources to cater for more needy students, at a deeper level, in the Tier 2 Program. Second, there are 40 units per grade (each lasting for 30 minutes), with a total of 120 units for the whole Tier 1 Program. Third, the program was developed by the Research Team by integrating existing research findings, programs, local adolescent needs, cultural characteristics and experiences (such as trial teaching) gained from the Experimental Implementation Phase of the Project. Fourth, relevant adolescent

developmental concerns (e.g., drug issues, sexuality, financial management, sense of responsibility, and life meaning) and adolescent developmental strengths (e.g., high level of concern for society and high proficiency in information technology) were incorporated in the program. Fifth, to cater for the needs of the students, the program is implemented by teachers and/or social workers.

The final attribute of the curriculum design is that the Project P.A.T.H.S. covers 15 positive youth developmental constructs which were identified from the existing successful positive youth development programs (Catalano, Berglund, Ryan, Lonczak & Hawkins, 2002). These constructs include: 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 fostering prosocial norms.

There are two implementation phases in this project – Experimental Implementation Phase and Full Implementation Phase. For the Experimental Implementation Phase (from 2005/06 to 2007/08 academic year), 52 secondary schools participated in the project with the objectives of accumulating experience in program implementation and familiarizing frontline workers with the program design and philosophy. In 2006/07 school year, the project was implemented on a full scale at Secondary 1 level. In 2007/08 school year, the project was implemented at Secondary 1 and Secondary 2 levels. In 2008/09 school year, the project was implemented at Secondary 1, Secondary 2 and Secondary 3 levels. In the present study, diaries written by students participating in the Secondary 1 Program (Experimental and Full Implementation Phases) and Secondary 2 Program (Experimental Implementation Phase) were collected to evaluate the Tier 1 Program.

The Project P.A.T.H.S. is a positive youth development program. In contrast to

mainstream approaches that focus on youth developmental problems, the field of positive youth development (PYD) focuses on the talents, strengths, interests, and future potentials in children and adolescents (Damon, 2004). According to Catalano, Berglund, Ryan, Lonczak and Hawkins (2002), there are several attributes of the positive youth development approach, including emphasis on integrated youth development (i.e., focusing on a range of youth developmental possibilities and problems) rather than dealing with a single youth problem, upholding the belief that "problem-free is not fully prepared", underscoring person-in-environment perspective, and focusing on developmental models on how young people grow, learn and change.

Although views differ in the strengths that should be developed in adolescents, many researchers suggested that building cognitive, academic, social and emotional competence is a fundamental task in adolescence (Graczyk et al., 2000). For example, Weissberg and O'Brien (2004) suggested that there are 5 core social-emotional competencies to be targeted in positive youth development programs: self-awareness, social awareness, self-management, relationship skills and responsible actions. Guerra and Williams (2003) similarly highlighted five core competencies for healthy youth development, including positive identity (positive self-concept, hopefulness, future goals), personal agency (self-efficacy, effective coping, locus of control, attributional style), self-regulation (affective, behavioral, and cognitive self-regulation, impulse control), social relationship skills (social problem solving skills, empathy, conflict resolution, capacity for intimacy), and prosocial system of beliefs (attitudes, norms, values, moral engagement).

The concept of positive youth development is closely linked to the construct of quality of life. Although there are different views on the definitions of quality of life, there is general agreement amongst researchers (e.g., Felce & Perry, 1995; Wallander, Schmitt & Koot, 2001) that the concept is a multi-dimensional one, including material well-being (finance, income, housing quality, transport), physical well-being (health, fitness, mobility, personal safety),

social well-being (personal relationships, community involvement), emotional well-being (positive affect, mental health, fulfillment, satisfaction, faith/belief, self-esteem), and productive well-being (competence, productivity). Obviously, positive youth development can be regarded as a cause, concomitant, component or consequence of quality of life, particularly with reference to social, emotional and productive well-being.

Roth, Brooks-Gunn, Murray and Foster (1998) pointed out that positive youth development programs are commonly regarded as "programs that provide opportunities and support to help youth gain the competencies and knowledge they need to meet the increasing challenges they will face as they mature" (p.423). Although many Western programs have been developed to promote positive development in children and adolescents, not all of them are successful in promoting adolescent development. For example, Catalano, Berglund, Ryan, Lonczak and Hawkins (2002) reviewed 77 programs on positive youth development. The review showed that only 25 programs were successful. Obviously (suggestion:Evidently), evaluation is an important issue to be considered in positive youth development programs.

Program evaluation is not a simple and straightforward task and there are many types and approaches of evaluation. In his discussion of the major strategies of evaluation, Patton (1997) outlined three basic types of evaluation: quantitative evaluation, qualitative evaluation, and utilization-focused evaluation. Ginsberg (2001) summarized the major forms of evaluation, including quantitative and qualitative approaches, cost-benefit analyses, satisfaction studies, needs assessment, single-subjects designs, experimental approaches and models, utilization-focused evaluation, empowerment evaluations, fraud and abuse detection, client satisfaction and journalistic evaluation. Using starting alphabets as the bases of classification, Patton (1987) suggested that there are more than 100 types of evaluation. In a comprehensive evaluation of the major evaluation paradigms, Patton (2002) pointed out that different attributes are associated with different evaluation paradigms.

As pointed out by Patton (1997), there are two main approaches in the field of

evaluation. For the quantitative/experimental paradigm, quantitative data are commonly collected through standardized and uniform procedures as well as fixed and controlled designs such as experiments with treatment and control groups. In addition, deductive hypothesis testing involving independent and dependent variables, linear and sequential modeling, pre-post focus on change, probabilistic and random sampling, statistical analysis, as well as generalizations are intrinsic to this paradigm which emphasizes objectivity in program evaluation and detachment of the evaluator from the program. For the qualitative/naturalistic paradigm, qualitative data such as narratives are collected through naturalistic inquiry, case studies, emergent and flexible designs and inductive analysis. In addition (suggestion: Furthermore), holistic contextual portrayal, systems perspective, interdependencies, dynamic and ongoing view of change, purposeful sampling of relevant cases focusing on uniqueness and diversity, thematic content analysis are intrinsic to this paradigm which emphasizes subjectivity in program evaluation and close relationship between the researcher and the program.

An examination of the evaluation literature shows that although there is a strong preference for the use of quantitative or experimental approach to evaluate the effectiveness of prevention programs for adolescents, the number of qualitative evaluation studies in this area is increasing (Shek, 2008). Moreover, regardless of the general impression that quantitative evaluation is the "mainstream" evaluation approach, there is a growing emphasis on the use of qualitative methods in evaluation. For example, Chatterji (2004) criticized the sole reliance on experimental methods and argued for the use of more diverse types of evaluation strategies, particularly in the context of education. Slayton and Llosa (2005) also remarked that "qualitative methods should be an essential part of large-scale program evaluations if program effectiveness is to be determined and understood" (p. 2543).

To understand the program effects of the Project P.A.T.H.S., several complementary program evaluation strategies including both quantitative and qualitative have been adopted.

These included objective outcome evaluation, subjective outcome evaluation, process evaluation, interim evaluation, and qualitative evaluation based on focus groups, in-depth interviews and case studies have been employed to evaluate the effectiveness of the project (e.g., Shek, 2006b; Shek, Ma, Lui & Lung, 2006; Shek & Ma, 2007). Through the adoption of different evaluation strategies, there is evidence supporting the effectiveness of the Tier 1 Program of the Project P.A.T.H.S. based on different types of data collected from different participants utilizing different methods (Shek, 2008). For example, both quantitative findings based on randomized group trial and qualitative findings from different studies provide evidence that the Project P.A.T.H.S. is effective in promoting the psychosocial competencies of the program participants (Shek, 2008).

To further understand the perceived program effects of the Tier 1 Program, evaluation findings based on weekly diaries written by students are reported in this paper. In the context of social sciences research, diary is commonly used in the clinical and educational contexts to chart human behavior. There are also studies using diaries or reflective journals as an evaluation strategy to examine program effects (e.g., Schmitz & Wiese, 2006). Furthermore, it is a common practice for teachers in Hong Kong to ask students to write weekly diaries or reflective journals to document their views on current events and their experiences. With the ease of collecting students' weekly diaries, the studies reported in this paper utilized students' weekly diaries as a research strategy to evaluate the perceived attributes and effectiveness of the Project P.A.T.H.S.. As there are both Experimental Implementation Phase and Full Implementation Phase in the Project P.A.T.H.S., it is reasonable to examine whether the findings are consistent across studies. As such, weekly diaries written by Secondary 1 students in the Experimental Implementation Phase, Secondary 1 students in the Full Implementation Phase, and Secondary 2 students in the Experimental Implementation Phase were collected and analyzed in Study 1, Study 2 and Study 3, respectively.

As a qualitative evaluation strategy, the question of "rigor" in weekly diaries is an

important question to be considered. Shek, Tang and Han (2005) suggested that 12 principles should be observed in a qualitative evaluation study. The first principle is that an explicit statement of the philosophical stand of the study should be given. The second principle is that the number and nature of the participants of the study should be justified. The third principle is that the data collection procedures should be clearly described. For the fourth and fifth principles, the biases and preoccupations of the researchers should be outlined and the steps taken to guard against biases should be described respectively. The sixth principle is that measures of reliability, such as inter-rater reliability and intra-rater reliability should be included. For the seventh and eighth principles, measures of triangulation in terms of researchers and data types and checking procedures (such as peer checking and member checking) should be included. The ninth principle is that the researcher should be conscious of the importance and development of audit trails. For the tenth and eleventh principles alternative explanations for the observed findings and explanations for negative evidence should be addressed. Finally, limitations of the study should be clearly stated. In the studies reported in this paper, the above principles were upheld as far as possible. In terms of philosophical orientation of the study, a general qualitative orientation utilizing qualitative principles (e.g., holistic emphasis and reliance on non-numerical raw data) was adopted.

<div align="center">Study 1</div>

Participants and Procedures

Among the 52 schools joining the Tier 1 Program (Secondary 1 level) in the Experimental Implementation Phase in 2005/2006, four schools were randomly selected to join this research study. The school teachers were asked to randomly invite some Secondary 1 students to write a journal in the form of weekly diary to reveal their perceptions and feelings after joining the Tier 1 Program or randomly select the submitted weekly diaries to the Research Team. The students were informed of the purpose of this study and confidentiality of the data. The consent of the participants was also sought. The total number of students'

weekly diaries received was 95, which was about 15% of the total Secondary 1 students in these four schools. As the cases were randomly selected and the response rate was 100%, the recruited sample could be regarded as respectable.

Data Analyses

The data were analyzed using general qualitative analyses techniques involving three steps (Miles & Huberman, 1994). First, relevant raw codes were developed for words, phrases and/or sentences that formed meaningful units at the raw responses level. Second, the codes were further combined to reflect higher-order attributes at the category of codes level. Third, the categories of codes were further analyzed to reveal the broader themes at the thematic level.

In the present qualitative analyses, as the author designed the program in the Project P.A.T.H.S., he was conscious of his own biases and expectations of the program to be effective. As such, the author was not directly involved in the first few steps of the data analyses. Both intra-rater and inter-rater reliabilities on the coding regarding the views on the program and instructors (Table 1) and perceived benefits of the program (Table 2) were calculated. For intra-rater reliability, the research assistants responsible for the original coding re-coded 20 randomly selected responses and the mean reliability percentage was computed. For inter-rater reliability, another person (e.g., a doctoral student or a research assistant with a master degree) not involved in the original coding re-coded 20 randomly selected responses without knowing the codes finalized by the author at the end of the scoring process.

Results and Discussion

As shown in Table 1, 146 meaningful units regarding students' perceptions of the Tier 1 Program and the instructors could be categorized into two categories (i.e., views on program and views on instructors). Overall, most of the responses regarding the students' perceptions of the program and instructors were positive. For the responses on the perceived benefits (Table 2), a total of 203 meaningful units were categorized into five categories (i.e., societal,

familial, interpersonal, personal levels of competence and others). Most of the respondents reported that they learnt personal competence, following by moral competence and virtues, cognitive competence, and emotional competence.

Several cases illustrate the positive perceptions of the informants:

Student A: "Although "P.A.T.H.S." is part of the Integrated Humanities subject, I think they are very different. P.A.T.H.S. is very interesting. The instructor always cracked jokes to create a relaxed atmosphere. Students also gave creative responses to make the class feel that it was a relaxed course. In contrast to other subjects where the classroom atmosphere is solemn, I think this course makes the students feel relaxed. As this course is relaxed and I can learn much knowledge from it, I like P.A.T.H.S. very much."

Student B: "This program enabled me to learn a lot of things which could be applied in my daily life. It helped me develop team spirit, enhance cooperation amongst group members and acquire independence. It was very easy to understand the content of every lesson. Although we were sometimes noisy, the instructors were very patient in teaching us and they would only teach when classmates were quiet. Because of this, I have learned self-discipline".

Student C: "After joining P.A.T.H.S., I learned how to deal with problems when I face adversity. I also learned much wisdom of life and have more self-understanding. Before the course, I did not have any dreams and goals. This course helped me develop my goals, understand my dreams and think about how to fulfill my dreams. When I faced difficulty in the past, I only cried. However, it is different now as I have learned different ways to solve my problems in class. Although these methods cannot totally solve my problems, I will not simply cry now."

Study 2

Participants and Procedures

There were 207 schools joining the Secondary 1 Program of the Full Implementation Phase of the Project P.A.T.H.S.. After completion of the Tier 1 Program, six schools were

randomly selected to join this research study, with the school teachers randomly invited some Secondary 1 students to write a journal in the form of weekly diary to reveal their perceptions and feelings after joining the Tier 1 Program. The students were informed of the purpose of this study as well as confidentiality and anonymity of the data collected. The students' willingness to join the study was respected and their consent of participation was sought. A total of 216 Secondary 1 students participated in the study. As the cases were randomly selected and the response rate was 100%, the sample size could be regarded as respectable.

In order to ensure consistency of the data collection procedures, a clear guideline on data collection was given to the teachers concerned. The students were asked to write a reflective journal with the title of "Participation in the Tier 1 Program of the Project P.A.T.H.S: Experiences and Feelings". The weekly diary was expected to be not less than 200 words in Chinese and it should be related to the students' experiences, feelings and comments in connection with their participation in the Tier 1 Program. The students could either complete it at home or during class time.

Results and Discussion

The data analyses procedures and principles were identical to those adopted in Study 1. Basically, the codes emerged from the content of the weekly diaries and they were developed with reference to the coding system devised in Study 1. As shown in Table 1, 369 meaningful units regarding the students' perceptions of the Tier 1 Program and the instructors could be categorized into two categories (i.e., views about program and views about instructors). In the aspect of "views about program", there were 319 responses which could be categorized into "overall impression", "program content", "learning process", and "other comments and suggestions". On the other hand, there were 50 responses of "views about instructors". In total, there were 84.6% positive responses regarding the students' perceptions of the program and instructors.

The perceived benefits of the program to the students are shown in Table 2. There were

752 meaningful units which could be categorized into five categories (i.e., societal, familial, interpersonal, personal levels of competence and others). Most of the respondents reported that they had learnt personal competence, following by "moral competence and virtues", "emotional competence", and "cognitive competence". Next, many respondents reported that they acquired interpersonal competence. Below are three cases illustrations on the perceived benefits of the program:

Student A: "When I first joined the program, I thought it was a program that did not make sense and it wasted our time. However, I gradually felt that the program was very useful – it increased my self-confidence and improved my relationships with friends, classmates and family members. As such, I have gradually liked this program. From the transition period from Primary 6 to Secondary 1, we need the concern of our family, friends and teachers. This program can exactly help us how to live. I remember that there is a lesson entitled "I was born with talents". I understand that everybody has talents. The most important point is how we view ourselves and discover our talents. Besides, as time is precious and important, we have to treasure our time".

Student B: "In the project, I could learn more about life from mutual sharing. I really learned a lot from the program. The most impressive lesson was "How to interact with other people", such as tolerance and practice which cannot be learned from the books. This program is not boring and I have interest in it. Life is a process. The sad thing is that we cannot live twice. The joyful thing is that we do not have to live twice as we know how to cherish time. After joining the program, I have become more optimistic and cheerful. Not only do I know how to control my emotions but also have more friends now.

Student C "I think the program is very meaningful. I have watched many video clips in the process. I was moved by the situations in the poor countries such as Africa. I hope I could send some clothes to them".

Study 3

Participants and Procedures

Among the 49 schools joining the Tier 1 Program (Secondary 2 level) of the Experimental Implementation Phase, four schools were randomly selected to join this research study. The school teachers were asked to randomly invite some Secondary 2 students to write a journal in the form of weekly diary to reveal their perceptions and feelings after joining the Tier 1 Program. The students were informed of the purpose and confidentiality of the research, and their consent of participation was sought. The students were asked to write a reflective journal with the title of "Participating in the Tier 1 Program of the Project P.A.T.H.S: Experiences and feelings". The reflection was expected not less than 200 words in Chinese and it should be related to the students' experiences, feelings and comments after joining the Tier 1 Program, though there were no specific requirements on how to write it. The students could either complete it at home or during class time. The total number of students' weekly diaries received was 260, which was about 36% of the total Secondary 2 students in these four schools. As the cases were randomly selected and the response rate was 100%, the sample size could be regarded as respectable.

Results and Discussion

The data analyses procedures and principles were identical to those adopted in Study 1 and Study 2. Basically, the codes emerged from the content of the weekly diaries and the coding system was developed with reference to those devised in Study 1 and Study 2. As shown in Table 1, 843 meaningful units regarding the students' perceptions of the Tier 1 Program and the instructors could be categorized into two categories (i.e., views about program and views about instructors). In the aspect of "views about program", there were 684 responses which could be categorized into "overall impression", "program content", "learning process", and "other comments and suggestions". Most of the respondents had positive views on the "overall impression". There were 159 responses of "views about instructors", which

were categorized into "overall impression" (e.g., "the instructor was pleasing"), "teaching performance" (e.g., "instructor's performance was excellent"), "teaching attitude" (e.g., "involved"), and "others". In total, there were 83.63% positive responses regarding the students' perceptions of the program and instructors.

The perceived benefits of the program to the students are shown in Table 2. There were a total of 635 meaningful units which could be categorized into five categories (i.e., societal, familial, interpersonal, personal levels of competence and others). Most of the respondents reported that they had learnt personal competence. In addition, many respondents reported that they had acquired interpersonal competence, which could be categorized into "general interpersonal competence" (e.g., "get along with others") and "specific interpersonal competence" (e.g., "respect").

Some narratives based on the informants revealed that the program was perceived positively by the program participants:

Student A: "In the process, the instructors were very dedicated in teaching us about the ways of life. I have learned how to respect others. Even though other people are not doing well, we should not blame them. Also, I have learned that none of us is perfect. Through the activities, I have also learned how to solve problems. Besides transmitting knowledge to us, the instructors also encouraged and supported us. In addition, when we encountered difficulties, we could share our problems with them who would then propose some suggestions and solutions to us."

Student B: "I am very fortunate to participate in the program as other schools may not have this opportunity. I have learned much knowledge and got valuable experiences from the program. I remember that in one lesson, the instructor asked those quiet students to answer the questions and I was one of them. When I answered the question, I was very nervous. However, when I answered more questions during class, I gained more confidence. The topics are specially designed for students and they are very practical. Most of the activities can raise

the self-esteem of the students. Students can really learn much from the program.

Student C: "When I first participated in the Project P.A.T.H.S., I really did not want to join as I did not understand the nature of the program. My thought at that time was even you gave me $1 million or $10 million, I still did not want to join it. However, after a few lessons, I began to like the program because the program enabled me to form close friendships with fellow classmates. I could also understand the past history of the instructor. Most important of all, the program has helped me reflect things in the past.

General Discussion

Utilizing weekly diaries written by the students, three studies are reported in this paper to lend further support to the program effectiveness of the Tier 1 Program of the Project P.A.T.H.S.. As far as the design and data collection of the studies are concerned, these three studies are qualitative studies, sharing the characteristics of naturalistic inquiry, inductive analysis, holistic perspective, qualitative data, unique case orientation and contextual sensitivity (Patton, 1990). Regarding data analyses, both qualitative (e.g., coding and categorizing) and quantitative (mainly counting) analyses were used. According to Miles and Huberman (1994), counting is a general tactic for generating meaning in qualitative research. Tashakkori and Teddlie (1998) also pointed out that mixed data analyses methods could be carried out for qualitative studies.

Based on the weekly diaries written by the students, the studies reported in this paper showed that the participants generally had positive perceptions of the Tier 1 Program of the Project P.A.T.H.S. and the instructors. The findings presented in Table 1 suggest that the participants had positive impression of the program and they appreciated the program content and learning process involved. Besides, Table 2 shows that the participants overwhelmingly regarded the program to be beneficial to the emotional, social and productive quality of life of the program participants. For example, a sizable number of responses were related to positive self-image after joining the program. In addition, many participants remarked that their

interpersonal competencies were enhanced after joining the program.

The above findings are generally consistent with previous research findings in the Experimental and Full Implementation Phases and they basically concur with both the quantitative and qualitative findings based on objective outcome, subjective outcome, qualitative and process evaluation findings (e.g., Shek, 2008). In short, the present findings provide additional support for the effectiveness of the Tier 1 Program of the Project P.A.T.H.S. in Hong Kong. From the perspective of triangulation, the existing evaluation findings suggest that the effectiveness of the Tier 1 Program of the Project P.A.T.H.S. is supported by evaluation data collected from different sources and by different strategies.

In the broader context of science, there are views arguing that replication is an important principle that should be used to evaluate scientific findings. For example, in an attempt to highlight the misinterpretations associated with statistical tests of significance, Cohen (1994) argued that social scientists "must finally rely, as has been done in all the older sciences, on replication" (p.997). Similar view was held by Shaver (1993) who suggested that "the notion of reproducibility leads directly to replication, widely agreed upon as a crucial element of science, but largely missing from reports of social science" (p.312). Shaver (1993) further argued that "editors should not only actively encourage the reporting of replications, but in many instances demand replication before results can be published" (p.312). Obviously, the scientific value of the evaluation findings based on weekly diaries written by the students can be substantially enhanced if we can produce evidence to show that the positive findings can be replicated across studies.

The present findings also underscore the utility of using weekly diaries to evaluate positive youth development programs. By asking the program participants to reflect their experiences, subjective perceptions and perceived benefits of the program can be properly understood in a less mechanical and artificial manner. Such an approach can help to capture the dynamic nature of the perceived qualities and effectiveness of the program. In addition to

its holistic emphasis and ease of data collection, Maxwell (2004) commented that "qualitative research is a rigorous means of investigating causality" (p. 3). Slayton and Llosa (2005) also remarked that "qualitative methods can confirm that it is actually the program that is responsible for the effect" (p. 2544). Obviously, the use of reflective journals enjoys the strengths of qualitative studies in assessing the changes in quality of life after joining the program.

Of course, credibility is a basic issue in qualitative research and it is no exception to the present studies. Following the principles of qualitative analyses (Shek, Tang & Han, 2005), the following attributes of the studies reported in this paper are highlighted: (a) a general qualitative orientation was adopted; (b) recruitment process for the participants and justifications for the number of participants are described; (c) details of the data collection are given; (d) the issues of biases and ideological preoccupation are addressed; (e) inter-rater and intra-rater reliabilities information is presented; and (f) the categorized data are kept in a systematic filing system in order to ensure that the findings are auditable.

Nevertheless, with reference to the arguments of Shek, Tang and Han (2005) that the authors should discuss the limitations of the qualitative evaluation studies conducted (Principle 12), several limitations of the studies reported in this paper should be noted. First, for students who are sensitive about the issue of "invasion of privacy" (Rothwell & Ghelipter, 2003), the use of weekly diary may be regarded as an obtrusive evaluation device. Second, as the diary method is verbal in nature, students who are not linguistically competent may write little to express their experiences. This problem may be particularly acute in schools admitting students with poor academic performance. This is also a problem in the Chinese culture where Chinese people may lack the language to describe their feelings. Third, as there was only one occasion through which the participants expressed their views, it was not possible to have dialogues with the informants to further understand some of their views. In particular, theoretical sampling approaches could be attempted in future studies.

Fourth, although the present findings are interpreted in terms of the positive program effects and experiences of the program participants, it should be noted that there are several alternative explanations. These alternative explanations should be considered as high proportions of responses were positive responses across the three studies. The first alternative explanation is that the students were afraid that they would be punished by the program implementers if they did not respond in the favorable direction. However, this possibility can be partially dismissed as the teachers did not mark the weekly diaries and the identity of the students was basically anonymous. The second alternative explanation is that the students consciously acted in a "nice" manner to help the workers to illustrate positive program effect. However, this alternative explanation could be partially dismissed because negative comments were recorded and students are normally encouraged to express their views in Hong Kong. The third alternative explanation is that the high proportion of positive responses observed in fact biased responses. However, this alternative explanation can also be dismissed because the schools and students were randomly selected.

Finally, although eleven principles proposed by Shek, Tang and Han (2005) were upheld in this study, peer checking and member checking (Principle 8) were not carried out in this study because of time and manpower constraints. Despite these limitations, this study provides pioneering qualitative evaluation findings over time that support the positive nature of the Project P.A.T.H.S. and its effectiveness in promoting holistic youth development and quality of life among Chinese adolescents in Hong Kong.

References

Catalano, R. F., Berglund, M. L., Ryan, J. A. M., Lonczak, H. S. & Hawkins, J. D. (2002). Positive youth development in the United States: Research findings on evaluations of positive youth development programs. *Prevention and Treatment, 5(1)*

Chatterji, M. (2004). Evidence on "what works": An argument for extended-term mixed method (ETMM) evaluation designs. *Educational Researcher*, *33*, 3-13

Cohen, J. (1994). The earth is round (p < .05). *American Psychologist*, *49*, 997-1003

Damon, W. (2004). What is positive youth development? *Annals of the American Academy of Political and Social Science*, *591*, 13-24

Felce, D. & Perry, J. (1995). Quality of life: Its definition and measurement. *Research in Developmental Disabilities*, *16*, 51-74

Ginsberg, L. H. (2001). *Social work evaluation: Principles and methods*. (Boston: Allyn and Bacon)

Graczyk, P. A., Weissberg, R. P., Payton, J. W., Elias, M. J., Greenberg, M. T. & Zins, J. E. (2000). Criteria for evaluating the quality of school-based social and emotional learning programs. (In R. Bar-On & J. D. A. Parker (Eds.), *The handbook of emotional intelligence* (pp.391-409). San Francisco, CA: Jossey-Bass.)

Guerra, N. G. & Williams, K. R. (2003). Implementation of school-based wellness centres. *Psychology in the Schools*, *40*, 473-487

Maxwell, J. A. (2004). Causal explanation, qualitative research, and scientific inquiry in Journal item: Education. *Educational Researcher*, *33*, 3-11

Miles, M. B. & Huberman, A. M. (1994). Qualitative data analysis. (Thousand Oaks, CA: Sage)

Patton, M. Q. (1987). How to use qualitative methods in evaluation. (Newbury Park, CA: Sage)

Patton, M. Q. (1990). Qualitative evaluation and research methods. (Newbury Park, CA:

Sage)

Patton, M. Q. (1997). Utilization-focused evaluation: The new century text. (Thousand Oaks, CA: Sage)

Patton, M. Q. (2002). Two decades of developments in qualitative inquiry: A personal, experiential perspective. *Qualitative Social Work, 1*, 261-283.

Roth, J., Brooks-Gunn, J., Murray, L. & Foster, W. (1998). Promoting healthy adolescents: Synthesis of youth development program evaluations. *Journal of Research on Adolescence*, *8*, 423-459

Rothwell, A. & Ghelipter, S. (2003). The developing manager: Reflective learning in undergraduate management education. *Reflective Practice*, *4*, 241-254

Schmitz, B. & Wiese, B. S. (2006). New perspectives for evaluation of training sessions in self-regulated learning: Time-series analyses of diary data. *Contemporary Educational Psychology*, *31*, 64-96

Shaver, J. P. (1993). What statistical significance testing is, and what it is not. *Journal of Experimental Education*, *61*, 293-316.

Shek, D. T. L. (2006b). Effectiveness of the Tier 1 Program of the Project P.A.T.H.S.: Preliminary objective and subjective outcome evaluation findings. *TheScientificWorldJOURNAL*, *6*, 1466-1474

Shek, D. T. L. (Ed.). (2008). Special Issue: Evaluation of Project P.A.T.H.S. in Hong Kong. *TheScientificWorldJOURNAL*, *8*, 1-94

Shek, D. T. L. & Ma, H. K. (2007). Subjective outcome evaluation of the Project P.A.T.H.S.: Findings based on the perspective of the program participants. *TheScientificWorldJOURNAL*, *7*, 47-55

Shek, D. T. L., Ma, H. K., Lui, J. H. Y. & Lung, D. W. M. (2006). Process evaluation of the Tier 1 Program of the Project P.A.T.H.S. *TheScientificWorldJOURNAL*, *1*, 300-309

Shek, D. T. L., Tang, V. & Han, X. Y. (2005). Quality of qualitative evaluation studies in the

social work literature: Evidence that constitutes a wakeup call. *Research on Social Work Practice*, *15*, 180-194

Slayton, J. & Llosa, L. (2005). The use of qualitative methods in large-scale evaluation: Improving the quality of the evaluation and the meaningfulness of the findings. *Teachers College Record*, *107*, 2543-2565

Tashakkori, A. & Teddlie, C. (1998). Mixed methodology: Combining qualitative and quantitative approaches. (Thousand Oaks, CA: Sage)

Wallander, J. L., Schmitt, M. & Koot, H. M. (2001). Quality of life measurement in children and adolescents: Issues, instruments and applications. *Journal of Clinical Psychology*, *57*, 571-585

Weissberg, R. P. & O' Brien, M. U. (2004). What works in school-based social and emotional learning programs for positive youth development. *Annals of the American Academy of Political and Social Science*, *591*, 86-97

Table 1: Views on the program and instructors in Study 1, Study 2 and Study 3

Views on the Program	**Reponses**	**Study 1**	**Study 2**	**Study 3**
Overall Impression	Positive (e.g., meaningful, excellent, meets adolescents' needs, better than other programs)	26	116	277
	Negative/neutral (e.g., mixed blessings)	3	1	18
	Total	29	117	295
Program Content	Positive (e.g., practical, creative, rich content, easily understood)	27	25	68
	Negative/neutral (e.g., dislike growth puzzles)	6	30	8
	Total	33	55	76
Learning Process	Positive (e.g., funny, interesting, happy, active participation)	34	70	107
	Negative/neutral (e.g., boring)	5	18	96
	Total	39	88	203
Other Comments	Positive/neutral comments	23	55	104
	Negative comments	0	4	6
Total		124	319	684
Views on the Instructors				
Overall Impression	Positive (e.g., positive style)	2	8	41
	Negative/neutral (e.g., disgusting)	3	1	3
	Total	5	9	44
Teaching Performance	Positive (e.g., excellent, prepared well)	9	28	51
	Negative/neutral	0	3	4
	Total	9	31	55
Teaching Attitude	Positive (e.g., enthusiastic)	5	7	40
	Negative/neutral	0	0	0
	Total	5	7	40
Other Comments	Non-negative comments	3	3	17
	Negative comments	0	0	3
Total		22	50	159
Grand Total		146	369	843

Note: Twenty coded raw descriptors were randomly selected for examining consistency in categorization of responses (views on program vs. instructors). Intra-rater reliability = 100% and inter-rater reliability = 85% in Study 1. Intra-rater reliability = 100% and inter-rater reliability = 85% in Study 2. Intra-rater reliability = 100% and inter-rater reliability = 85% in Study 3.

Table 2: Perceived benefits of the Tier 1 Program in Study 1, Study 2 and Study 3

Psychosocial Domains	Sub-Domains	Study 1	Study 2	Study 3
Societal Domains	Social norms	3	12	3
	Social responsibilities and knowledge about public affairs	11	33	8
	Total	14	45	11
Familial Level	Family relationship	5	22	12
Interpersonal Level	General interpersonal competence (e.g., making friends, getting along with others)	42	162	147
	Specific interpersonal competence (e.g., respect, empathy)	6	63	28
	Total	48	225	175
Personal Level	Cherishing life	2	24	53
	Reflection	8	14	37
	Cognitive competence	12	46	54
	Ways to face adversity	1	17	33
	Positive self-image	32	124	88
	Emotional competence	11	57	28
	Goal setting	6	23	29
	Moral competence and virtues	14	99	47
	Total	86	404	369
Others	Learn much knowledge	14	26	31
	Learn something useful	13	18	0
	Interested in learning	1	4	10
	Others	22	8	27
	Total	50	56	68
Grand Total		203	752	635

Note: Twenty coded raw descriptors were randomly selected for examining consistency in categorization of responses (i.e. changes on different levels). Intra-rater reliability = 100% and inter-rater reliability = 85% in Study 1. Intra-rater reliability = 100% and inter-rater reliability = 85% in Study 2. Intra-rater reliability = 100% and inter-rater reliability = 85% in Study 3.