

## **Research Evaluation and Scale Development of the Wholistic Education for the Leadership Learning Project (WELL) I**

This report summarizes overall research evaluation of the Wholistic Education for Leadership Learning Project - Stage One (WELL I). The evaluation focused the overall effectiveness of programmes as indicated by both process and outcome measures from participating students and teachers. Quantitative and qualitative research methodologies and evaluation techniques were adopted to gather information. Different modes of data indicated the overall effectiveness of the WELL I project.

In addition, two standardized scales have been developed as a parallel component of research. The two scales have been constructed and validated in local secondary school settings, representing initial attempts in the contextualization and development of psychometrically sound instrument for use in Chinese youth education and leadership training.

### **Background**

Time frame of the research evaluation process and scale development of WELL I last for one year, consisted of framework formulation, sampling, measurement, research design, data collection, data analysis, and report writing. From September 1998 to September 1999, a total of 40 individual programmes have been evaluated, with different themes, nature of activity and means of delivery. In addition, two standardized scales, namely Resilience and Empathy, have been constructed amongst adolescent samples. These two instruments measured independent dimensions of the 12 wholistic leadership qualities, has been validated for local use.

In order to nurture different sub-components of the 12 leadership qualities, six types of educational modules have been designed in WELL I programmes, namely the Wholistic leadership camp, Residential theme camp, Day camp, Sequential programme, Train-the-trainer programme, and Underachiever & Resiliency programme. Individual programmes focused on a single theme or combination of multiple themes to foster a wholistic approach leadership education.

These educational programmes were delivered by a number of front-line workers with a wide spectrum of specialties such as training, media production, arts and performance, personal growth & development, adventure activities, communication & group dynamics, and youth education etc.

The evaluation analyses have focused on the overall effectiveness of the project as well as different components according to individual educational modules. As the scale development is concerned, major properties and limitations of the two developed scales will be discussed. Detailed literature review, sampling and validation procedure, and psychometrics of these new scales have been included in separate full reports.

### **Method**

For the WELL I evaluation component, procedure has been intended a wider approach involving formative, summative, and integrative evaluation techniques. Standardized questionnaires (SQ) were widely used to attain mainly quantitative

responses regarding satisfaction levels on different objectives, contents and formats in various programmes. As an understanding of process views and subjective impact of programme on participants, focus group discussions were used to acquire more dynamic and in-depth information for more continuous and extensive programmes.

In addition, a pre-and-post test evaluation design has been used to pilot for a more rigorous testing of different levels of change of participants upon the completion of an extensive programme (i.e. White Dolphin Action 99) which fostered changes in resilience related factors.

A parallel component of research evaluation was the scale development of the overall WELL I research. Scale development procedure included the formulation of conceptual and measurement models of the two constructs, constructions of measurement items and scenarios, piloting of assessment instrument, data collection, analyses and report writing.

The evaluation instrument administered mainly by means of a standardized questionnaire (SQ) developed at the early phase of the WELL project. After meetings and discussions with individual programme developers involved in the project, format and content of the SQ was tailored made for the use of various WELL I programmes.

The instrument consists of two sections, namely part one and part two. The first part is an administrative record of individual programmes, filled out by programme staff, included information such as focus of leadership factors, type of educational module, type and duration of activities or programmes, and number of participants.

The second part is the process evaluation form, filled out by participants, included information concerning subjective satisfaction ratings on the over and specific programme contents, format and means of delivery and staff performance. A few open-ended questions taped on the individual objectives of joining the programme, most unforgettable and impressive programme contents, and any necessary improvement for the progamme. Information on gender, age and educational levels were also included in the end of SQ to provide demographic information of participants.

As administration procedure is concerned, both SQ and focus group discussion would normally be administered at the end of each programme as part of concluding activities. For sequential progrmmes, several SQs may be used at different stages of the programme to collect information concerning specific objectives and progress of participants.

Evaluation information was available for analyses to examine the specific process, overall effectiveness and ultimate outcome of individual and overall programmes of the WELL I project.

## **Results**

### **(A) Research Evaluation**

Overall, 40 individual programmes have been delivered to serve a total of 2,165

participants comprising 1,698 students, 467 teachers and social workers. Students participated in various training programme included the wholistic leadership camp, theme course, day camp and sequential course. The majority of students (i.e. 1,668) came from main stream secondary schools, which were main targets for training of various wholistic leadership qualities. Thirty students from lower banding schools, commonly labeled as underachiever, participated in the pilot sequential programme (i.e. White Dolphin Action 99), an extensive programme which fostered changes in resilience-related factors. On the other hand, teachers completed the train-the-trainer programmes, which are on-job training for designing and implementing student leadership learning modules at school.

Results from the quantitative and qualitative evaluations indicated the overall satisfactory level of various programmes. Specific aspects of programmes according to various educational modules will be described in the following.

### **Wholistic leadership camp**

The wholistic leadership camp is an intensive comprehensive programme, lasting for one week, fostering various experiences and skills related to nurturing wholistic leadership qualities. In order for students to grasp a whole range of leadership concepts and skills, the programme targeted more senior form secondary students, with high quality leaderships and being quick learners of their schools.

Results from standardized questionnaire (SQ) indicated more than 86% of participants was satisfied with the overall programme content, 41% reported a 'very good' response and 12% rated an 'average' response. Qualitative data showed that 30% of participants was impressed by reflection experience ('山上獨處之新鮮體驗及反思'), 35% enjoyed the 'team building and group experience' ('要完成一件事, 必須互相合作和彼此用心去對待'), and 27% finding the 'adventurous training experience' helpful ('能夠玩到石牆和繩網這些特別的活動, 在當中能夠使我面對和克服困難').

As format and delivery of programme is concerned, more than 80% of participants was satisfied, 26% rated a 'very good' response and 18% gave an 'average' response. The majority (92%) of participants was satisfied with performance of programme staffs, with half of responses giving a 'very good' rating. Qualitative data showed that 27% of participants appreciated the performance of staffs. Positive comments were given concerning the training style, friendliness, sincerity and concern for participants. ('PANDA 的訓練方式', '全部導師認真, 投入', '突破工作人員的充足及富心思的準備以及親切, 關心的表現').

As a whole, 79% of participants agreed that the programme was good in meeting individual objectives of participants and less than 16% gave an 'average' response and less than 4% showed dissatisfaction regarding this aspect. Qualitative data showed that about 60% of participants stated that leadership learning was their main objective ('希望可以學習怎樣成為一位全面的領袖') and 20% intended to have personal growth ('認識自己和改進自己, 發掘自己的才能').

Regarding possible improvement for the programme, most participants were concerned about the timing and arrangement of programmes. Qualitative data showed that timing was too tight, too many programme activities, preference of interactive programme activities to lectures etc. A number of suggestions were given included as possible improvement of overall programmes (‘時間太緊迫’, ‘活動進行過於密集, 學生沒有適當的時間去休息’, ‘在講坐裡有一點悶, 比較重資料性, 希望可以加些樂趣’).

### **Residential theme camp**

Residential theme camp included a total number of ten programmes, varied from two days overnight camp to three days & two night camps, focus on one main leadership qualities. The programmes targeted different forms of secondary students, nurturing different focused qualities, suitable for particular stages of their school lives. A number of themes such as emotional intelligence, resilience to adversity, creativity & media education, creativity & IT education, and multiple intelligence were intended as important elements to be addressed.

Results from standardized questionnaire (SQ) indicated more than 71% of participants was satisfied with the overall programme content, 17% reported a ‘very good’ response and 27% rated an ‘average’ response. Qualitative data showed that most participants was impressed by a number of activities delivered by individual programmes, such as ‘adventurous training experience’, ‘reflection experience’, ‘video taking’ and ‘Expo 2020 project’ etc. (‘EQ 歷奇挑戰站’, ‘攀石’, ‘山上獨處的一刻, 令我真在認識到自己’, ‘可以自己親自拍一輯電影’, ‘Expo 2020 中各組 present 其 project, 能看見他們所負出的努力, 而且參考其他人的創作意念’, ‘蒙眼上山’).

As format and delivery of programme is concerned, more than 63% of participants was satisfied, 15% rated a ‘very good’ response and 31% gave an ‘average’ response. The majority (80%) of participants was satisfied with performance of programme staffs, with half of responses giving a ‘very good’ rating.

As a whole, 60% of participants agreed that the programme was good in meeting individual objectives of participants and less than 34% gave an ‘average’ response and less than 6% showed dissatisfaction regarding this aspect. Qualitative data showed that participants most intended to learn leadership, self-growth, EQ, media production, new technology, know new friends and to play (‘領袖培訓’, 增加自信和 EQ’, ‘學建媒介製作的技巧, 啟發創意’, ‘認識創意與科技的關係’, ‘認識新朋友’, ‘貪玩’).

Regarding possible improvement for the programme, 31% of participants was concerned about the timing, 42% on programme design and facility. Various suggestions included time was too short, some activities were boring, content too difficult and quality of food etc. (‘時間太短’, ‘看電影很無聊’, ‘說話太長’, ‘有些講坐很悶’, 講坐層次高了些, 對那未曾接觸媒介的同學來說, 顯得深奧’, ‘飯太硬’).

### **Day camp**

Day camp included a total number of eight programmes, targeted both junior and senior forms of secondary students, nurturing different focused qualities in relatively short exposures. Various qualities such as emotional intelligence, team building and group dynamics, adventurous mind, resilience to adversity, media & creativity, and multiple intelligence were selected as focuses of programmes.

Results from standardized questionnaire (SQ) indicated more than 74% of participants was satisfied with the overall programme content, 14% reported a 'very good' response and 21% rated an 'average' response. Qualitative data showed that most participants was impressed by a number of activities as delivered by individual programmes, such as 'adventurous training experience', 'public speaking', 'musical show' and 'serving experience etc. ('無懼柱', '攀石', '演講', '用七巧板砌東西', '用樂器來講故事', '音樂工作坊', '認識如何服侍人和欣賞人').

As format and delivery of programme is concerned, more than 68% of participants was satisfied, 18% rated a 'very good' response and 29% gave an 'average' response. The majority (83%) of participants was satisfied with performance of programme staffs, with 30% of responses giving a 'very good' rating.

As a whole, 61% of participants agreed that the programme was good in meeting individual objectives of participants and less than 37% gave an 'average' response and less than 3% showed dissatisfaction regarding this aspect. ('提高EQ', '增強自信心', '訓練口才、思維', '希望成為一個出色的領袖', '希望能學習與別人相處, 懂得與人溝通', '不用上課, 有得玩', '發掘自己的潛能、創造力', '可以幫助有需要的人').

Regarding possible improvement for the programme, almost 19% of participants was concerned about the timing, 31% on programme design and facility and 13% of comments was related to resource and food quality of camp site. Various suggestions included time was too short, some activities were boring, and quality of food etc. ('延長時間, 例如兩日一夜', '講解時較悶', '部份設施如爬石遊戲', '太少體育用品', '食物質素').

### **Sequential programmes**

Sequential programmes are series of integrative and extensive programme, consisting of a variety of continuous programmes, fostering various leadership skills and experiences in stepwise manners. A number of five programmes targeted both junior and senior form students from a number of secondary schools. Various programme approached have been adopted such as the adventure-based and integrative counselling programme in the White Dolphin 99 (白海豚行動 99), sequential workshops of the learning of media production in the Reporter Workshops (小記者工作坊), a combination of workshops, camp and experiential learning in the Dream Wall Programme (2001 夢工牆), and an experiential approach of musical and drama performance in the My Story Express (我有我故事).

Results from standardized questionnaire (SQ) indicated more than 73% of participants was satisfied with the overall programme contents, 12% reported a 'very good'

response and 26% rated an 'average' response. Qualitative data showed that 51% of the White Dolphin 99 participants was impressed by the reflection and hiking experience ('自己一個人在山上寫信給父母'), and they also enjoyed the variety of funny and exciting activities ('有很多刺激活動', '很好玩'). Over thirty percent of participants enjoyed the things learned being helpful or the team cooperation ('學到東西太多', '我隊有合作精神').

For the Dream Wall Programme, 77% of participants was satisfied with the overall programme contents, 13% reported a 'very good' response and 23% rated an 'average' response. Qualitative data showed that 16% of participants was impressed by the 'adventurous training experience' ('攀石'), 44% mentioned the painting of wall artwork was most impressive ('畫牆畫'), and 19% most enjoyed the team cooperation ('大家的合作精神').

For the My Story Express Programme, 79% of participants was satisfied with the overall programme contents, 14% reported a 'very good' response and 21% rated an 'average' response. Qualitative data showed that 39% of participants reported that they have achieved learning the concepts and skills of script and lyric writing for drama ('嘗試填詞', '學了如何寫劇本') and 29% mentioned that the programme was good enjoyment or personally helpful ('不錯的娛樂', '比以前較自信').

For the Reporter Workshops, 47% of participants was satisfied with the overall programme contents. Qualitative data showed that 67% of participants reported that they were impressed by the experience of interviewing for news ('和別人做採訪') and 58% enjoyed the experience of learning to be a news reporter ('有機會讓我們參與報刊的製作', '有機會嘗試去做記者').

Taking together five programmes, more than 67% of participants was satisfied with format and delivery of programme, 13% rated a 'very good' response and 31% gave an 'average' response. The majority (85%) of participants was satisfied with performance of programme staffs, with half of responses giving a 'very good' rating. Qualitative data showed that participants appreciated the performance of staffs. Positive comments were given concerning the training and professional style, helpfulness, concern for participants ('導師的協助', '設備完善, 導師專業、細心', '導師能和同學融為一體').

As a whole, 62% of participants agreed that programmes were good in meeting individual objectives of participants and less than 35% gave an 'average' response and less than 4% showed dissatisfaction regarding this aspect. Regarding possible improvement for the programme, most participants were concerned about the timing and arrangement of programmes. Qualitative data showed that timing was too tight, too many programme activities, preference of interactive programme activities to lectures etc. A number of suggestions were given included as possible improvement of overall programmes ('時間短, 不能深入了解', '太多理論, 沒實踐', '課程教授方式, 生動為佳, 可用遊戲', '更多表演機會').

### **Train-the-trainer programme**

Train-the-trainer programmes included a total number of nine programmes, targeted

teachers of secondary schools, enhance understanding of a number of leadership qualities and experience in the practical implementation of wholistic leadership education through both seminar and experiential learning. Various emphasis included an integrative counselling approach to cultivate students' resilience, youth and creativity, cultural heritage and leadership education, group camp counselling approach and on-site school training.

Results from standardized questionnaire (SQ) indicated almost 90% of participants was satisfied with the overall programme content, 25% reported a 'very good' response and 10% rated an 'average' response. Qualitative data showed that most participants was impressed by a number of activities as delivered by individual programmes, such as 'content of video show', 'adventurous activities', 'creative teaching demonstration', 'seminar on Chinese culture' and 'game and group activities' etc. ('青苗行動的影帶', '歷奇輔導', '遊戲活動', '創意自由我開始', '中國文化的堅忍, 及對學生的影響', '攀石', '運用個人想像力表現生命的看法', ').

As format and delivery of programme is concerned, more than 87% of participants was satisfied, 28% rated a 'very good' response and 12% gave an 'average' response. The majority (94%) of participants was satisfied with performance of programme staffs, with 37% of responses giving a 'very good' rating.

As a whole, 80% of participants agreed that the programme was good in meeting individual objectives of participants and 18% gave an 'average' response and less than 2% showed dissatisfaction regarding this aspect. ('提高EQ', '增強自信心', '訓練口才、思維', '希望成為一個出色的領袖', '希望能學習與別人相處, 懂得與人溝通', '不用上課, 有得玩', '發掘自己的潛能、創造力', '可以幫助有需要的人').

Regarding possible improvement for the programme, almost 50% of participants was concerned about the timing, programme design and facility and 18% of comments was related to staffs. Various suggestions included time was too short, some activities were boring, content too difficult and quality of food etc.

## **Discussion**

Both overall and specific results indicated the success of the WELL I project. The original design of the leadership learning project aims to foster a wholistic approach of nurturing various leadership qualities of youth, targeting students with different training needs. Based on numerous programme evaluation data of WELL I, both students and teachers participants were very satisfied with various aspects of programmes, such as overall content, format and delivery, and staff performance. As individual objectives were concerned, the majority of participants perceived that programme objectives largely matched with their expectations. A significant portion of participants placed leadership learning as an important aim of joining various WELL programmes.

Specific comparisons participants' ratings of programmes showed that the wholistic leadership training camp scored the highest rating of satisfaction, indicating students preferred a more eclectic and comprehensive approach in delivering various

leadership concepts and skills. Because various leadership factors are relatively independent and yet inter-related concepts, students may grasp the whole picture of essential elements of a good leader. Theme camp and day camp are relatively focused in the amount and format of information being delivered; however, may be deficient in explaining global picture. The wholistic approach of leadership learning not only explain the value of individual quality in personal character building, but also various ways different qualities may complement each other. However, the high intensity of information and the tight schedule of programme activities may be physically demanding to students.

In order for students to grasp a number of inter-related leadership concepts and skills in a stepwise manner, the design of sequential programmes allowed concepts that build-up in earlier part of programme served as foundations for forthcoming skills and concepts. However, very careful coordination and planning would be needed to highlight the connections between the more theoretical and practical components of the programme. In this respect, the evaluation process also need to match the overall well with the programme design, which place very much a dynamic challenge.

The complexity of a number of integrative activities means that both qualitative and quantitative information would be beneficial to the evaluation. More experience has been acquired through these evaluation processes, especially the acquisition of qualitative information by the focused-group procedure.

Lastly, further improvement of outcome evaluation techniques will be needed in order to quantify important outcomes of programme delivery. For instance, more rigorous design of pre-and-post outcome measures, with the possibility of introducing the experimental design, in which outcome evaluation include both experimental and control groups in the evaluation process.

## **(B) Scale Development**

### **Background**

Since there are lacking local Chinese scales in measuring adolescence character building, it is intended that scientifically valid instruments will be developed on the whole range of leadership related qualities. Two psychometric scales, namely empathy and resilience were developed in WELL I, using two independent student samples from a number of secondary schools. These two dimensions are important dimensions amongst the 12 wholistic factors of the various leadership qualities.

Recent preventive programmes targeting youth developmental risk factors, focus a number of resilience-related as intervening factors such as students' competence, belonging, usefulness, potency and optimism (Sagor, 1996). Subsequent studies suggested that social support, self-esteem, problem-solving ability and purpose in life are four inter-related factors in determining the outcome of a resilient person (Werner, 1989; Whisman and Kwon, 1993; Prister and Clum, 1993; McMilian and Reed, 1994). Resilience is defined as 'the set of attributes that provides people with the strength and fortitude to confront the overwhelming obstacles they are bound to face in life' (Sagor, 1996, p.38). Resilience is also regarded to be a track record of successful adaptation after exposing to stressful life events.

On the other hand, empathy is another sector of interpersonal competence involving both affective response and cognitive perspective taking. Empathy is defined as 'the vicarious experiencing and accurate understanding of another person's feeling and point of view' (Yau, 1999, p.5). The central roles of empathy in determining altruistic behaviours and social competence make the assessment of empathy essential in character building and leadership training.

### **Scale construction and further development**

In the resilience scale, seven sets of questionnaires were used to construct a resilient outcome classification scheme, aims to identify individuals who have experienced adversities but are able to maintain psychological well-being and positive means of managing life stress.

Findings showed that resilient individuals obtained higher scores on a number of the four mediating factors, such as self-esteem, social support, problem-solving skills and purpose-in-life. Predictive validity of high resilient group vs. low resilient group using the four mediating factors, showed that purpose of life come out as the most powerful predictor. As the classification of high resilient group vs. depressive group is concerned, both purpose-in-life and problem-solving are significant predictors of the outcome.

Findings showed that the existing classification scheme remains at its preliminary stage and further modification of item contents will allow the construction of a more conceptually coherent and face valid measure of resilience. In short, these results enrich the understanding of resilience-related factors in affecting the outcome of resilient response, the existing work form a good start for the development of a multi-dimensional full scale of resilience.

As the empathy is concerned, an Empathy Scale (EMS) has been developed as a scenario-based self-report measure, and a western-developed scale, the Interpersonal Reactivity Index (IRI) has been validated locally.

The finalized version of EMS consists of 10 scenario-based statements, each with a pair of response options, representing either empathetic or non-empathetic responses. Concerning the psychometric property, Cronbach's alpha of 0.67 was reported, indicating a fairly internally consistent instrument. By using scenario as empathy-provoking stimulus and forced-choice dichotomous responses, the EMS is shown to tap actual empathy as a reactive construct than merely a self-perceived disposition. It is capable of predicting concern for the needy under the same situation, but it is unclear how much this may be generalized to other situations. Moreover, the external predictive validity of EMS is still largely unexplored. Though the EMS has its limitations, it compensates for the inadequacy of traditional measures such as the IRI. To establish the value of the EMS, it is recommended that the scale may be used in further testing with youth developmental programmes such as social skills training and social services. Moreover, comprehensive assessment of empathy may be undertaken in future research, so that the EMS may be used complementarily to the IRI for better inferences, judgements, and decisions.