

Designing Assessment Tools for Community Programmes



Dr Daniel SY Ho & Dr Xiaonan Yu
School of Public Health
University of Hong Kong

6 Dec 2010

1

Learning objectives

1. Have an overview of assessment
2. Recognize different components of a questionnaire at pre-test, post-test and follow-ups
3. Use process and outcome evaluation
4. Understand concepts of reliability, validity and sensitivity of assessment tools
5. Discuss strengths and limitations of using existing scales and designing new tools
6. Discuss important issues of designing assessment tools
7. Set clear and effective assessment items

2

Examples of community programmes: FAMILY Project: Intervention Studies

- Goal: Promotion of Family Health, Happiness, and Harmony
- Preventive interventions
- Efficient and sustainable
 - Use community interventionists
 - Minimize the number of sessions
 - Use group format
- Randomized Controlled Trial (RCT)
 - HKU Arm
 - NGO Arm
 - Control Arm
- Pre- and post-intervention design
 - Before the intervention (T1)
 - Immediately after the intervention (T2)
 - At three months after the intervention (T3)
- Evidence-generating

3

FAMILY: Effective Parenting Programme

愛+人：「有教·無慮」家庭和諧計劃

- **Partner:** Caritas-Hong Kong (香港明愛)
- **Target group:** Mothers of children in P1-P3
- **District:** Tin Shui Wai
- **Objectives:**
 - Improve parents' emotion management
 - Improve positive parenting
 - Enhance family health, happiness, and harmony
- **Research Design:** Three arm RCT (50 per arm)
 - Caritas arm
 - HKU arm
 - Assessment-only control arm
- **Length:** 4 x 2 hour sessions

4

FAMILY: Harmony@Home

愛 + 人：「家多·和諧」計劃

- **Partner:** Hong Kong Family Welfare Society (FWS) (香港家庭福利會)
- **Target group:** Parents of child aged 10 - 13
- **District:** Tseung Kwan O
- **Objective:**
 - Decrease parent-child conflict
 - Enhance parent-child relationship
- **Research Design:** Three arm RCT (50 per arm)
 - FWS programme
 - HKU programme
 - Waitlist control
- **Length:** 4 x 2-3 hour sessions

5

1. Overview of assessment

6

Evaluation

- Evaluation is the process of assessing what has been achieved (whether the specified goals, objectives and targets have been met) and how it has been achieved.
(Simnett, 1995)
- Programme evaluation is the use of social research procedures to systematically investigate the effectiveness of social intervention programs.
(Rossi, Lipsey, & Freeman, 2004)

7

Challenges in programme evaluation

- How to assess programme implementation?
- How to assess the outcome changes?
- How to identify desirable outcomes?
- How to capture relatively small effects in programmes?
- How to tailor the item wordings and response options to the target participants?

8

The importance of assessment

- Generate evidence about the effectiveness of interventions
- Quantify baseline status
- Determine if objectives have been met
- Identify what is working and what is not working in programme
- Justify the use of resources
- Help Improve our practice by learning our successes and mistakes
- Assist future planning
- Inform other researchers and practitioners to avoid making the same mistakes

Help you to become a better service provider!

9

Characteristics of good assessment

- Based on programme framework, theory and objectives
- Focused on desirable outcomes
- Integral to programme design
- Cost-effective
- Have clear instructions
- Have reliable and valid measurements
- Provide useful information
- Quantitative
- Comparable to other studies and programmes
- Have feedback and reflection

10

Sources of data in programme evaluation

- Self-administered surveys
- Achievement tests
- Medical record reviews
- Direct observations
- Interviews
- Physical examinations
- Clinical scenarios
- Performance tests
- Literature review

(Fink, 2005)

11

Consider assessment at different stages

- **Intervention design**
 - Think about how to match intervention and assessment
- **Trial group**
 - Test whether participants have difficulty understanding the items
 - Observe participants' responses
 - Get feedback from participants
 - Estimate time participants spent on the questionnaire
- **Pilot study**
 - Find out problems in assessment
 - Improve intervention and assessment based on preliminary results
- **Main study**
 - Generate evidence about the effectiveness of programme
 - Capture what is changed as the result of intervention
 - Review the strengths and limitations of assessment tools

Assessment is on-going throughout the programme!

12

2. Components of a questionnaire

13

Components of a questionnaire

	Pre-test	Post-test	Follow-up
Informed consent	✓		
Instructions	✓	✓	✓
Personal identifiers (complete or partial) for matching	✓	✓	✓
Socio-demographic characteristics	✓		
Process evaluation items		✓	
Outcome evaluation items	✓	✓	✓

14

Informed consent

- **Key elements**
 - Research purpose; evaluation procedures; estimated time
 - Use of data collected and by whom; data safety; right to access information
 - Confidentiality; right to withdraw without consequences;
 - Participant's signature; date; contact person and number for evaluation
- **Children participants**
 - Design separate informed consent forms for parents
 - Have verbal forms for children

Informed consent protects all parties: participants, evaluator and the institution

15

Instructions

- **Criteria of good instructions**
 - Precise
 - Concise
 - Confirmed by interventionists
- **Example: FWS Harmony@home**
 - 回答此問卷時，請誠實地作答。答案沒有對錯之分，請不要花費太多的時間去考慮。
 - 作答時，請以你現年其中一名10至13歲的子女為對象。
 - 請在格內以“✓”代表你所選擇的答案。

16

Personal identifiers

- For matching between anonymous pre- and post-tests
- Failure to match reduces effective N
- Explain reason for collecting these data
- Full info for less sensitive identifiers
 - Date of birth (YYYY, MM, DD)
- Partial info for more sensitive identifiers
 - Telephone numbers (e.g. last 4 digits)
 - HKID numbers (e.g. first 3)
- Avoid identifiers that are prone to change
 - E.g. # of children / siblings; marital status; occupation; education
- May use multiple identifiers
 - Depends on uniqueness and N
- Use available identifiers
 - School code, grade, class (change after promotion), sex

17

Socio-demographic characteristics

- Demography
 - Sex, age, place of birth, living district
- Family/household related
 - Marital status; who lives in household; # of children; maid?
- Socioeconomic status
 - Less sensitive: education level, occupation group, housing type, house size
 - More sensitive: personal income, household income, CSSA, asset worth
 - For children: parental education/occupation, perceived family affluence, # of bedrooms

18

3. Process evaluation and outcome evaluation

19

• Process evaluation

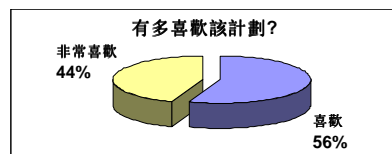
- The need for the programme
- Design of the programme
- Programme implementation and service delivery
- How participants think about the programme?
 - How enjoyable? How useful? How likely to recommend?

• Why process evaluation is important?

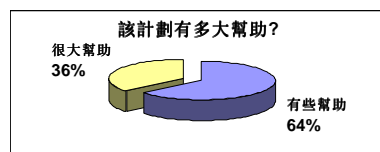
- Monitors progress: pace? efficient? cost-effective?
- Monitors programme quality & participant feedback
- Helps explain the observed effects

Process evaluation is nearly always useful (Fink, 2005) ²⁰

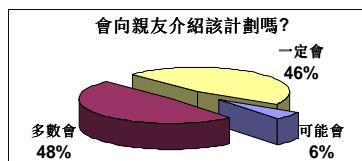
Customer satisfaction



(1=完全不喜歡, 5=十分喜歡)



(1=完全沒有幫助, 5=很大幫助)



(1=絕對不會, 5=絕對會)

(支援組, n=36)

21

Qualitative feedback from participants

支援組的參加者認為計劃內容十分有用，能夠幫助他們有效地改善情緒及促進家庭關係。

– 「我肯聽孩子的說話，自己又情緒好左。覺得家中安靜，大家都沒有發脾氣。」— 孫太

– 「我丈夫變得更投入家庭生活。過往，他總是無視我和孩子，現在，家庭氣氛的轉變令大家的關係變得更和諧、更快樂。」— 劉太

22

Outcome evaluation

- **Outcome evaluation**
 - What has been achieved?
 - Benefits or changes due to programme
- **The importance of outcome evaluation**
 - Generate evidence
 - Inform effectiveness, goal attainment and unintended outcomes
 - Is critical in quality improvement of programme

23

Time effect in outcome evaluation

- **Immediate: Pre-test & post-test**
 - Changes in knowledge, skills, attitudes, intention, behaviours?
- **Short-term: 3-6 months after the (completion of) programme**
 - Behaviour change
- **Long-term: One year or longer after the programme**
 - Changes in clinical health or quality of life

24

Examples of outcome items

- **Knowledge**
 - Proper sitting posture can prevent back pain
 - Emotion control is important in maintaining good relationship with family members
- **Ability and skills**
 - How to perform breathing exercises in order to relax
 - How to manage emotion in frustration
- **Attitudes**
 - Smoking is harmful to health
 - Parents with psychiatric problems should not raise their children

25

• Intentions

- I intend to control negative emotions when my child misbehaves

• Behaviours

- Talk to family and friends when feeling sad
- Cigarettes smoked per day
- Frequency of exercise

• Health status

- Illness in past 2 weeks
- Perceived health

26

Design assessment based on intervention framework and objectives

- **Use programme goal, objectives, and intervention model to guide assessment**
 - Is outcome measurable?
 - Is outcome changeable?
 - Is assessment matching your programme goal and objectives?
 - Measurement for each intervention component?

27

Identify primary and secondary outcomes

• Primary outcome

- First benefits or changes experienced and the ones most closely related to programme outputs
- Occurs during or at the end of programme

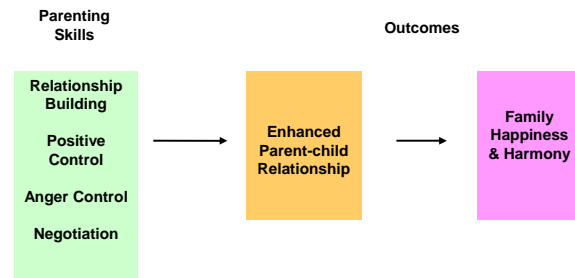
• Secondary outcomes

- Long-term consequences
 - usually the ultimate goals

Are your outcomes clearly defined, measurable, and achievable?

28

FWS – Positive parenting Intervention Model



- **Primary outcome**
 - Satisfaction with relationship with child
- **Secondary outcome**
 - Family Harmony
 - Happiness

29

Exercise

- Discuss the primary and secondary outcomes in the following programme
 - E.g.: We are conducting a programme to promote effective adaptation of new arrivals. They will attend an intervention programme which provides them information about education, medical care, housing, employment and community resources in Hong Kong. What are the outcomes in this programme?

30

- **Primary outcome**
 - Knowledge about education, medical care, housing, employment, and community resources
 - 我清楚香港幼稚園及中小學的人學申請辦法
 - 我瞭解學費減免申請資格
 - 我明白政府門診預約方法
 - 我瞭解公屋申請資格及方法
 - 我瞭解香港搵工途徑
- **Secondary outcome**
 - Perceived social support (我能夠找到合適的機構協助)
 - Community involvement (我會參與社區活動)
 - Sense of belonging (我是香港的一份子)

31

4. Reliability, validity and sensitivity

32

Reliability

- Repeatability / consistency of measurement
- E.g. Emotion management items
 - Deal with the situation calmly when your child misbehaves
 - Manage your anger when your child misbehaves
 - Talk to your family or friends when you feel stressed
 - Try to find some ways to relax yourself in order to stabilize your emotions
- Internal consistency (items measuring same concept)
 - Cronbach's alpha: 0.72
- Stability
 - Test-retest reliability

33

Validity

- The degree to which a test assesses what it intends to measure
 - Factor structure
 - Principal component analysis results of Caritas emotion management items:
One factor, 56% of variance explained
 - Concurrent validity
 - Correlation with other emotion management measures
 - Predictive validity
 - The extent to which a measure forecasts future performance

34

Validity

Ask the research question from different perspectives

- Relationship with children
 - 對於你和子女的關係，你有多滿意？
極之不滿意 / 非常不滿意 / 稍微不滿意 / 尚算滿意 / 非常滿意 / 極之滿意
- Conflict with children
 - 在過去兩星期內，你和子女發生多少次衝突 (衝突在這包括意見不合,不服從, 吵鬧, 打架等)？
0次 / 1次 / 2次 / 3次 / 4次 / 5次或以上
 - 在過去兩星期內，你和子女衝突的程度有多強烈(衝突在這包括意見不合,不服從, 吵鬧, 打架等)？
沒有 / 輕微 / 稍微強烈 / 強烈 / 嚴重

35

Sensitivity

- The smallest amount the scale can measure
- Sensitivity to time change
 - Problem: "Do things with your child that you both enjoy"
 - Never
 - Seldom
 - Sometimes
 - Often
 - Always
 - Revision: "In the past two weeks, how frequently did you do something nice for your child to enhance your relationship?"
 - Not at all (Zero times a week)
 - Rarely (1-2 time a week)
 - Sometimes (3-4 time a week)
 - Often (5-6 times a week)
 - Almost always (7+ times a week)

36

Modify existing scales in your study

- Change the content to make them suitable for your programme components
- Tailor the items to make them understandable to your target population
- Modify the responding method to make them sensitive to time change
- Change the item order to group positive items and negative items separately

41

Considerations for designing new tools

- Have achievable and measurable objectives
- Target age group
- Education background
- Time available
- Environment of assessment
- Length of questionnaire

42

FWS FAMILY: Harmony@home

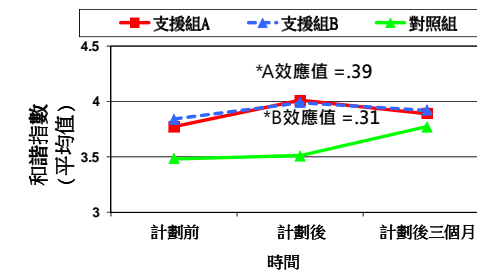
Example: Using self-developed items

- Family Harmony
 - 以下是一系列有關閣下家庭的句子。各人對家庭的定義有所不同，一般而言，家庭成員包括有血緣關係、姻親關係、同住，或感情關係密切的人。請根據你個人對家庭的定義，於每一題目選出最合適的答案。
 - 我家各人相處融洽。
 - 一家人住在同一屋簷下很開心。
 - 總括來說，我滿意我的家庭。
 - 相比其他家庭，我家各人的關係很密切。
 - 我家日常相處和睦。
 - 我的家庭是和諧的。
 - 對於各家庭成員，我們的家庭運作良好。
 - 我的家庭是一個快樂的地方。
- 1=非常不同意 2=不同意 3=中立 4=同意 5=非常同意

43

FWS FAMILY: Harmony@home

- Harmony results in FWS pilot study
 - Participants in both Arms A and B reported enhanced levels of Harmony post-intervention vs. the Control group.



* $p < .05$ · 具有統計學意義

注釋1. 所有比較均控制了基線值

注釋2. 效應值(Effect Size), 小 = .10, 中 = .25, 大 = .40

44

6. Important issues of designing assessment tools

45

FWS FAMILY: Harmony@home

• Objective assessment

- Participants were asked to rate the *frequency of performing a specific behaviour* in the past two weeks.
- They were assessed with the same questions
 - At baseline (T1),
 - After the intervention (T2)
 - At 3-month follow-up (T3).

– “在過去兩星期，你有多經常與子女談論他/她感興趣的話題，即使你對該話題不感興趣？”

- 1=從來沒有
- 2=甚少
- 3=有時
- 4=頗多
- 5=經常

46

FWS FAMILY: Harmony@home

• Subjective assessment

- Participants are asked to indicate *the amount of perceived change*, if any, at the end of the intervention (T2) and at follow-up (T3), compared to the time before they join the programme (T1)

– “與參與此課程之前作比較，你有改變與孩子討論他/她有興趣的話題的次數嗎？”

- 1=沒有改變
- 2=稍微增加
- 3=有些增加
- 4=增加頗多
- 5=增加很多

47

FWS FAMILY: Harmony@home

Table 1. Changes of targeted parenting behaviours from T1 to T3 in FWS study

Parenting behaviours	Change in frequency of behaviour from T1 to T3		Perceived change from T1 to T3	
	p value	Effect size	p value	Effect size
Discuss with child the topics that interest him/her	.853	.001	.001	.127
Give reasonable consequences	.651	.003	.001	.119
Give clear expectations when child misbehaves	.037	.053	.018	.067
Negotiate good behaviours with child	.300	.013	.226	.018
Stay calm when child argues back	.454	.007	.051	.046

Effect size (Eta-squared), small=0.01, medium=0.06, large=0.14

48

Pros and cons of subjective measures

- **Pros**
 - Direct and quick
 - Overcome low sensitivity of some objective measures
 - Cover a wider scope (e.g. perceived health) than objective measures
 - Avoid the need to objectively specifying some abstract concepts
 - Can be asked by nonprofessional staff
- **Cons**
 - Subjective feeling of change may not indicate actual change
 - There may be a concern of social desirability
 - Might not be appropriate for every construct measured
- **Conclusions**
 - Subjective measures tend to produce more favourable results
 - Use both subjective and objective measures in assessment

49

Common problems of item wordings

- **Double-barreled questions**
 - Problem
 - Express affection by hugging and kissing your child
 - Revision
 - Express affection by hugging your child
 - Express affection by kissing your child
 - Problem:
 - Are you happy with your relationship with family members and friends?
 - Revision:
 - Are you happy with your relationship with family members?
 - Are you happy with your relationship with friends?
- **Double negative**
 - Problem
 - I do not think that she is not a good teacher
 - Revision
 - I think that she is a good teacher
 - Problem
 - My family is not an unhappy family.
 - Revision:
 - My family is a happy family.

50

• Confusing and ambiguous

- Problem
 - Do you negotiate good behaviour with your child?
- Revision
 - When it comes to something my child doesn't want to do, I discuss my child's ideas with him/her and come to a mutually agreeable solution
- Problem
 - To what extent do you agree or disagree that you are happy?

Strongly disagree/Disagree/Agree/Strongly agree
- Revision
 - I am happy

Strongly disagree/Disagree/Agree/Strongly agree

The item wording should be clear, simple, and specific.
Avoid complex terms, jargons and unclear words

51

• Difficulty recalling and reporting past behaviours

- Problem: 當子女不聽話時，責罵或嘮叨他/她
- Revision: 在過去兩星期內，當子女不聽話時，責罵或嘮叨他/她
- Revision: 一般來說，你平均每天有多少次當子女不聽話時，責罵或嘮叨他/她

Keep recall simple and related to recent events

• Ceiling effect

- The highest option is not high enough, so most subjects selected the highest option
 - E.g. Elderly QoL scale used in adolescents

• Floor effect

- The lowest option is not low enough; most chose the lowest

52

Common problems of response options

- **Too few**
 - Yes/no
- **Too many**
 - Agree Strongly
 - Agree Moderately
 - Agree Slightly
 - Tend to Disagree
 - Neutral
 - Tend to Agree
 - Disagree Slightly
 - Disagree Moderately
 - Disagree Strongly

53

• Not mutually exclusive

- Problem: From which one of these sources did you get information about this programme?
 - In Integrated Family Service Center
 - In school
 - A friend
 - A social worker
- Revision: Where did you get information about this programme?
 - In Integrated Family Service Center
 - In school
 - In a street booth
- Revision: Who referred you to attend this programme?
 - A school teacher
 - A social worker
 - A friend or relative

54

• Insufficient options

- “與參與此課程之前作比較 (Control group—與填寫第一份問卷時作比較)，你有改變與孩子討論他 / 她有興趣的話題的次數嗎？”

– Problem:

- 沒有改變
- 稍微增加
- 有些增加
- 增加頗多
- 增加很多

– Revision:

- 減少很多
- 減少一些
- 少許減少
- 中立
- 少許增加
- 增加一些
- 增加很多

55

• Inappropriate ranges

- Problem: 家庭每月總收入
 - < \$10,000
 - \$20,000 – \$30,000
 - \$30,000 – \$40,000
 - \$40,000 – \$50,000
 - \$50,000或以上
- Problem: 最近一年以來，你看過幾次醫生？
 - ≤ 10
 - 11-20
 - 21次或以上
- Revision: 家庭每月總收入
 - < \$2,000
 - \$2,000 – \$5,999
 - \$6,000 – \$9,999
 - \$10,000 – \$19,999
 - \$20,000 – \$29,999
 - \$30,000或以上
- Revision: 最近12個月內，你看過幾次醫生？
 - 0 – 4
 - 5 – 9
 - 10 – 14
 - 15 – 19
 - 20次或以上

56

Different wordings and options may yield different results

- How old are you? ___age
- What is your current age? ___Years
- When were you born? __Year born
- What is your date of birth? __Day__Month__Year
- Which of the following age category describes you:
 - 35 or younger
 - 36-65
 - 66 or older

(Dillman, 2000)

57

Closed vs open questions

- **Closed question**
 - Did you like the programme?
 - From strongly disagree to strongly agree
 - Do you have a good relationship with your parents?
 - Yes/no
- **Open-ended question**
 - What do you think of this programme?
 - How would you describe your relationship with your parents?
 - Strengths
 - Is informative, learn about unexpected topics or issues
 - Limitations
 - Require more work/effort, time-consuming

58

Considerations for using Likert scales

- **Avoid central tendency bias**
 - Asians are more likely to select midpoint and avoid extreme responses than Americans (Lee, Jones, Mineyama, & Zhang, 2002)
 - Take out the middle choice of “neither agree nor disagree” or “neutral”
- **Optimal length of a Likert scale**
 - Not more than 7 points
- **Minimize the variation of points in one questionnaire**
 - Use the same number of points on as many items as possible
- **Place items with same response categories together**

59

Have a pilot test before the study!

- Sample size
- A live test with procedures intended to be used in the main study
- Make use of preliminary evidence to refine your questions and format

Pilot test in the field is a critical component of questionnaire design.
(Iraossi, 2006)

60

Checklist during pilot test

- Do the respondents understand the objective of the survey?
- Do the respondents feel comfortable answering the questions?
- Is the wording of the survey clear?
- Is the time reference clear to the respondents?
- Are the answer choices compatible with the respondents' experience in the matter?
- Do any of the items require them to think too long or hard before responding? If so, which ones?
- Which items produce irritation, embarrassment, or confusion?
- Do any of the questions generate response bias? If so, which ones?
- Do the answers collected reflect what you want in regards to the purpose of the survey?
- Is there enough diversity in the answers received?
- Is the survey too long?
- According to your test audience, have any other important issues been overlooked?

(Iraossi, 2006) 61

Recommended readings

- Fink, A. (2005). Evaluation fundamentals: Insights into the outcomes, effectiveness, and quality of health programs. 2nd edition. Thousand Oaks: Sage.
- Dillman, D.A. (2000). Mail and internet surveys: The tailored design method (2nd ed.). New York: John Wiley & Sons, Inc.
- Iraossi, G. (2006). The power of survey design: A user's guide for managing surveys, interpreting results, and influencing respondents. Washington, D.C.: The World Bank.

62

7. Set clear and effective assessment items (Practical session)