

## 6.6. Monitoring and Evaluation

Monitoring and evaluation of The Pillowcase Project has, for the most part, followed a similar pattern across six of the participating jurisdictions (there was no evaluation of the Vietnam initiative).

First, a pre-test and post-test of children's level of knowledge of disaster preparedness has been applied, data from the pre-test being used as a baseline against which the results of a largely or fully identical post-test could be judged. This has taken the form of a short quiz or questionnaire (USA, Hong Kong, Peru and, still at the design stage, Mexico). In the case of Australia there was no pre-quiz for students but a post-session multiple-choice quiz was administered.



*Completing Post-Session Quiz, USA*

Second, post-session quizzes have incorporated one or more questions to ascertain student perceptions of whether they feel more emergency-prepared or confident as a result of the intervention (USA, Australia, Hong Kong, Peru and, again still at the design stage, Mexico). In the case of Australia, this has taken the form of a question appended to the post-session multiple-choice quiz. In the case of the United Kingdom *Before* and *After* questionnaires use a ten-point scale to ask about student confidence in preparing for an emergency, whether it is a good thing to prepare for an emergency and whether the respondent feels friends and family would support them in an emergency. Added to the *After* questionnaire are checking boxes to identify who the respondent intends to share their learning with, a question about the most important thing learned from The Pillowcase Project session and space to draw a picture of the respondent's favorite part of the session.

Third, feedback on the impacts of the session has been garnered from significant adults, i.e. presenters, assistants, host teachers, parents and guardians. This has taken the shape of a presenter/assistant evaluation form (USA, Australia, Hong Kong), a host teacher session feedback/evaluation form (USA, Australia, Hong Kong), an impact feedback form for teachers (USA, Australia), a hardcopy or online survey or questionnaire on impacts for parents (USA, Australia), a teacher survey seeking details of program implementation, feedback on the level and quality of student engagement and on what worked best in the session (United Kingdom) and a feedback form for 'service users', i.e.

host institutions (Hong Kong). Mexico is currently developing mechanisms for gathering parental impressions of the program.

There are some examples of data collection methods that go beyond the above-delineated pattern. In Hong Kong a sample of students were interviewed about their responses to the post-session questionnaire. Hong Kong has also regularly used post-class focus group sessions involving volunteers, teaching assistants and host teachers to review the content, tutor performance and other aspects of the lesson. Similarly, the Peruvian Pillowcase Project team has held local evaluation workshops involving volunteers, school principals, parents and others. The Hong Kong team experimented with using a revision test with students some two months after the Project lesson but deemed the sample – drawn from one school - to be too small and unrepresentative to be valid.

We said earlier (sub-section 6.3) that an effective and meaningful evaluation of The Pillowcase Project has to be set against a clear understanding of intended learning outcomes or objectives. To this point the evaluation of the Project has only been loosely based upon stated learning outcomes, in part because they have not been fully set out. In this regard, we fully endorse the US team's intention to pursue 'a more intentional focus on priority learning objectives' in subsequent evaluations, and the British intention 'to measure the learning objectives and how the educational outcomes are met'. The task is one of determining the knowledge, skills and attitudinal learning outcomes intended by the Project, and then shaping the evaluation instruments and modalities so they are fit for purpose in finely assessing achievement against those intended outcomes (while also capturing unintended or unexpected outcomes).

So far, data collection from students has particularly focused upon measuring knowledge acquisition and improved confidence in being able to prepare for and cope with emergencies. Beyond perceptions of changes in level of confidence, wider attitudinal shift has not been measured. That could be remedied by having students react to a series of statements designed to evince a wide spectrum of attitudinal positions. *Figure 5* below offers a sample pre- and post-attitudinal questionnaire based upon the attitudinal learning outcomes proposed in *Table 1* (pp.68-70). If taken up, 'blind', i.e. off-topic, statements should be randomly inserted while a box for explanation of answers can be added after each statement as a means of garnering useful qualitative data. Likewise, skills development has not been monitored although it could be by observing, for example, how students perform in contrived pre- and post-session situations. More or less missing, too, is the use of longitudinal evaluation instruments to measure whether and to what degree new learning and attitudes resulting from Project session(s) hold over time. At the moment, most data is gathered during a session or in the fairly immediate aftermath of a session, the evaluation thus giving a snapshot of immediate session impact that may be diluted or reinforced over time.

### **Figure 5: Pre- and Post-Attitudinal Questionnaire**

	Strongly Agree	Agree	Don't Know/Not Sure	Disagree	Strongly Disagree
As children, we don't have to think about risks and emergencies. It's up to adults to do that for us					
I feel sure I am prepared if something dangerous - a fire, a flood, a really bad storm - should happen					
I don't really need to have an emergency kit ready at all times. People will look after me if we are faced with some threat					
I feel shy about sharing with others worries that I have about risks and threats to my and my family's safety					
I get upset about the dangers that people in other parts of the world face from natural hazards					
There are fire officers, police, doctors and nurses, and soldiers whose job it is to protect us. It's not so important that my family, neighbors and local people make their own preparations for disasters					
I believe that people in my community should act together so we are ready to face any danger					
As a young person, I can play an active part in helping make others safe from dangers					

The absence of longitudinal data is particularly problematic in terms of assessing whether the Share dimension of the learning framework has been achieved. In the case of a few countries, notably the USA and Australia, a survey, available online or in hard copy, seeks information from parents on what learning students passed on at home, what they and the family have done in consequence and what further disaster preparedness plans they have. In the case of Australia, returns are analyzed in the evaluation report to the Global Disaster Preparedness Center. Such parental surveys can make a potentially valuable contribution to understanding what happens and what is being achieved under the Share dimension of The Pillowcase Project but missing is an input from the children themselves. We would suggest that The Pillowcase Project Disaster Preparedness Advocacy/Action Rainbow typology (see p.76) might be used to find out what students have actually done by having them note down what they have done under each of the six levels of agency or leadership. This might be done some six to eight weeks after the Project lesson(s), even periodically thereafter, to understand the sharing and advocacy initiatives of students. It might be accompanied, as suggested earlier (p.67) by a 'show and tell' class session in

which students share and discuss what they have done and learnt. Another idea, also touched on earlier, would be to invite parents and children to attend a disaster-preparedness 'moot' where families share what steps they have taken in the wake of the Project. Either suggested event should be used not simply as a learning occasion but also to gather evaluative data on sharing.

Another useful longitudinal evaluation instrument that also provides a good learning reinforcement opportunity and that merits occasional repetition is the use of 'snap groups' to observe and assess the degree to which students have internalized disaster preparedness learning.

### **Snap Groups**

Students form small groups of three or four and are told that there is a sudden emergency described on the card that will be handed out. Groups are informed they have two minutes to note down all the things they need to have done and to do right now to reduce the chance of that emergency seriously harming them. After calling 'Stop!' the facilitator asks each group to report back and then asks other groups to respond and critique what they have heard. A series of cards are talked through. A short report on the quality of student disaster preparedness thinking as manifested in the session might be forwarded to the respective Red Cross team as updating data on the impact of the Project lesson(s). Three examples of cards are given below.<sup>1</sup>

#### ***In the classroom***

*In the classroom you are sitting at your desk during the math lesson when the classroom begins to shake violently. Windows rattle and books fall off the shelves. What do you do?*

#### ***At home***

*You and your family are having dinner when you hear on the radio a hurricane warning for your area. What do you do?*

#### ***Fire next door***

*In the next apartment you hear someone shout 'Fire!' and you smell smoke. What do you do?*

But there is a range of longer-term, repeatable evaluation interventions that could be employed, including:

---

<sup>1</sup> Activity taken from: Selby, D., & Kagawa, F. (2014). *Disaster Risk Reduction Education Toolkit*. St Michael: Caribbean Disaster Emergency Management Agency. 154-5, 170.  
[http://www.cdema.org/joomdocs/CDEMA\\_DRR\\_Edu\\_Toolkit\\_MAY\\_22\\_2015.pdf](http://www.cdema.org/joomdocs/CDEMA_DRR_Edu_Toolkit_MAY_22_2015.pdf)

- Repeat of evaluation instruments used after The Pillowcase Project lesson(s) at some six to eight weeks' distance (similar to the Hong Kong 'revision tests')
- Occasional short and random 'spot check' individual interviews reviewing the Project experience and checking disaster-preparedness knowledge levels
- Repeated focus group interviews, i.e. re-running the Hong Kong post-class focus group and Peru local evaluation workshop format several weeks after the Project lesson(s)
- Student self-assessments in which individual students write reflections on what they have learnt guided by some simple questions such as 'What I now know about disaster preparedness', 'What I feel I don't know', 'What questions and concerns I still have about being disaster-prepared'
- Disaster preparedness learning portfolios, students collecting together all their work and writing on disaster preparedness - something that can be drawn upon periodically to identify the vitality of what has been learnt and pinpoint what learning needs there still are.

**Recommendation 15:** *Having determined in detail The Pillowcase Project learning outcomes, monitoring and evaluation mechanisms should be determined that measure those outcomes (or, at least, those outcomes held to be of primary importance). The realization of a mix of knowledge, skills and attitudinal learning outcomes should be measured and ways found for longitudinal, not just immediate, measurement. It is particularly important to close the gap in evaluation of the learning effectiveness of the Share dimension of the Project framework.*

Generally speaking, there is a case for greater rigor in Project monitoring and evaluation. Key areas needing attention are as follows:

- The administration of pre-test questionnaires and quizzes during the early stages of the intervention and when some introductory segments of the program have already been experienced. This decreases the reliability of the test returns because the thinking of students might already have been influenced (i.e. 'contamination')
- The administration of pre- and post-tests in the presence of the staff member, volunteer or teacher who is leading the session. The presence of a person perceived as an authority figure may skew responses, so constituting duress, this being especially the case with questions aimed at eliciting response to the quality of the session(s)
- The tendency to over-rely on quantifiable data. As an Australian Pillowcase Project document puts it, 'it's easy to capture and assess', but richer and more nuanced understanding comes from the mix and interplay of quantitative and qualitative evidence.
- It seems that qualitative data, where it is collected, is not being analyzed but simply used to 'decorate' evaluations. While data open to quantitative analysis is graphed and analyzed, verbatim quotes drawn from qualitative data tend to be displayed in occasional quotation boxes with no justification of selection of the quotation sample and no thematic organization of emerging themes and trends in the data.

- Overall, there is a strong case for greater interfacing or triangulation of different data sets to determine whether each is telling the same story or whether conflicts and inconsistencies exist between the different kinds of data (and what the explanation for that might be).

All jurisdictions, save the USA - where evaluation data has been fed into general phase reports - have published dedicated evaluation reports that are ultimately formative in purpose, i.e. to suggest future amendments and improvements to program, including the evaluation processes to be used.

***Recommendation 16:*** *There should be greater rigor applied to evaluating the impact and outcomes of Project interventions with thoroughgoing triangulation of different data sets and with consequent reduced reliance on quantitative data. In evaluation reports the evaluation methodology needs to be clearly set out and the interplay of qualitative and quantitative data, and all the problems that throws up, analyzed (and seen to be analyzed). This can only strengthen the validity and hence promotion of The Pillowcase Project.*