

*RESEARCH AND PRACTICE NOTE /
NOTE SUR LA RECHERCHE ET LES MÉTHODES*

THE ESSENTIAL SKILLS SERIES IN
EVALUATION: ASSESSING THE VALIDITY
OF THE ESS PARTICIPANT WORKSHOP
EVALUATION QUESTIONNAIRE

Christian Dagenais
Université de Montréal

Luc Dargis-Damphousse
Université du Québec à Montréal

Julie Dutil
Centre for Liaison on Intervention and Prevention in the Psychosocial Area
Montreal, Quebec

Abstract: Since 2003, the Essential Skills Series training program developed by the Canadian Evaluation Society has been offered to more than 15 groups in the province of Québec. The evaluations of these workshops were based on the participants' reactions collected by a Participant Feedback Questionnaire. This article describes the process used to assess the structure of the questionnaire and document its psychometric properties in order to determine whether it covered all subjects addressed by the training program and the extent to which it measured the items it was intended to measure. The results suggest that the questionnaire is effective in measuring participant responses to all the relevant components of the training program. This procedure may interest professional evaluators who want to ensure that questionnaires used to evaluate training programs are suitable for their intended purpose.

Résumé : Le programme de formation développé par la Société canadienne d'évaluation, la Série des compétences essentielles en évaluation de programmes, a été offert à plus de 15 groupes au Québec depuis 2003. Les évaluations de ces ateliers étaient basées sur les réactions des participants recueillies à l'aide d'une fiche d'évaluation de leurs réactions. Cet article présente le processus ayant servi à évaluer la structure et les qualités psychométriques du questionnaire, de façon à déterminer s'il adresse tous les sujets

Corresponding author: Christian Dagenais, Université de Montréal, C.P. 6128
succ. Centre-ville, Montréal, QC, Canada H3C 3J7; <christian.dagenais@
umontreal.ca>

du programme de formation et à quel point il mesure ce qu'il doit mesurer. Les résultats montrent que le questionnaire mesure efficacement les réactions des participants à tous les éléments pertinents du programme de formation. Cette procédure de validation pourrait intéresser les évaluateurs professionnels qui souhaitent s'assurer que les questionnaires qu'ils utilisent pour évaluer des activités de formation sont appropriés.

BACKGROUND

The Canadian Evaluation Society (CES), under the direction of its Professional Development Committee, supported the development of the Essential Skills Series (ESS) in order to meet the needs of managers who assign or manage program evaluations and professionals who wish to acquire basic evaluation skills and expertise. The first version of the ESS was finalized in May 1999.

This training program is designed for all newcomers to the field of program evaluation, as well as managers responsible for program evaluation and anyone wishing to learn the basics of program evaluation. The training objectives are (a) to promote greater knowledge of program evaluation concepts, procedures, and standards of practice; (b) to apply this knowledge to the practice of program evaluation; and (c) to situate the role of program evaluation in the context of program planning and development.

The series comprises four workshops:

1. Understanding program evaluation (definitions, concepts, models, professional ethics)
2. Planning an evaluation (steps in the evaluation process, needs assessment, feasibility study)
3. Evaluating program process (tools to evaluate and improve program performance)
4. Evaluating outcomes (conducting evaluations based on program outcomes).

The training program is not confined to a single approach to the issue. Instead, it covers a range of basic concepts associated with this discipline. Participants learn how to apply various techniques in order to plan and conduct program evaluations.

The program was initially developed in English by the CES. In August 2003, the *Centre de liaison sur l'intervention et la prévention*

psychosociales (CLIPP) was commissioned by the *Société québécoise d'évaluation de programme* (SQÉP) to translate and adapt the series workshops into French and to conduct a pilot project to test the French-language version of the training program, at least once, in Québec. The agreement stipulated that the CLIPP would maintain the content of the program, but adapt its form to the Québec context. Accordingly, changes were made to the presentation format, examples were added, and the vocabulary and some terms were adapted and approved by members of the SQÉP Training and Exchange Committee.

PROCESS USED TO MEASURE PARTICIPANT REACTIONS

The CLIPP pretested the training program by delivering each of the four workshops to six groups of participants. Based on their feedback, additional adjustments were made to the format. The updated training program was offered on numerous occasions over subsequent years.

Between December 2003 and September 2010, the CLIPP offered the French version of Workshops 1 and 2 of the training program to 247 individuals (17 groups), and Workshops 3 and 4 to 222 individuals (16 groups) in Montreal, Quebec City, and Ottawa. Groups ranged in size from 11 to 24 participants.

As agreed with the SQÉP, the CLIPP evaluated participant reactions after each training session and produced a report for each group. These evaluations were used to monitor the training and adjust it, as required, to meet the participants' needs. At the end of each of the four training days, participants were asked to complete a Participant Feedback Questionnaire (PFQ). These questionnaires were used to measure participants' perceptions with regard to the achievement of their objectives, the novelty of the concepts conveyed, their assimilation of these concepts, the difficulty of the workshops, and the relevance of the workshops to their work. The organization of the workshops was evaluated as well (e.g., content, material, overall workshop quality), as were the quality of the facilitator and his or her delivery methods. Lastly, participants were invited to provide qualitative feedback on any of these items.

During the first workshop, participants also completed a participant profile. It included questions regarding their experience in the field of program evaluation, prior related training, education, profes-

sional sector, employer, and reasons for participating in the training program.

Overall, the results of the PFQ analysis were very positive (Dutil & Dagenais, 2008). More than 77.5% of the participants (Workshops 1 and 2: $n = 170$, and Workshops 3 and 4: $n = 149$) reported that they had met their objective by taking part in the training workshops. More than 90% of participants were satisfied with the various aspects of workshop organization (organization of the material, workshop content, audiovisual aids, and overall workshop quality). The level of difficulty was considered suitable by 75% of participants. Results differed significantly, however, when the groups were heterogeneous, that is to say when they were composed of participants from very different organizations. This can be explained by the fact that it was difficult to provide examples applicable to all participants of these groups. More than 72% of participants stated that the content would be useful for their work. Lastly, participants were virtually unanimous in their appreciation of the presentation method and facilitator (satisfaction rate of 96% or higher).

To ensure that all relevant information regarding participant reactions was gathered, the feedback questionnaire was adjusted twice, first in 2003 and again in 2007. After numerous years of use, it has now become possible and necessary to test the validity of the questionnaire in order to provide an objective assessment of its psychometric properties.

VALIDATION PROCEDURE

An assessment of the PFQ's validity was undertaken in 2010. The assessment procedure was inspired by the "target model" (Chiocchio, 1996a, 1996b). This model was chosen for two reasons: (a) it uses scientific criteria generally accepted by the research community, and (b) it represents a pragmatic approach shared by several program evaluators. This model integrates components of psychometrics in a consistent conceptual framework. It aims to successively evaluate the usefulness, validity, and reliability of an instrument, in light of the anticipated use of the instrument (Chiocchio, 1996a). The usefulness of an instrument is the cornerstone of this model and involves matching its intended use with the information it provides: in the present case, it was intended to obtain feedback from participants of a training program to improve the program. This component encompasses the other two components. Validity refers to the question-

naire's relevance to the concept to be measured, in other words, how it measures what it purports to measure (Hogan, 2003). The reliability of a questionnaire (Hogan, 2003), reflects the extent to which its measurement error level provides sufficiently reliable information for the anticipated use. All of these three components interact. Any errors arising in one of them will have impacts on the other two.

A study to assess the validity of the participant feedback questionnaire, inspired by Hogan's model (2003), was conducted in 2010.¹ It described the structure of the questionnaire and documented its psychometric properties in order to determine whether the questionnaire covered all subjects addressed by the training program and to assess the extent to which it measured the items it was intended to measure. The feedback questionnaire documents participant reaction (degree of satisfaction and perceptions) to each of the four ESS workshops. The responses gathered provide information for improving the training program and enabling an assessment of overall participant satisfaction.

Structure of the Questionnaire

For each of the workshops, the questionnaire measures six aspects (Table 1) of participant reaction. Almost half of the aspects are quantitative. The remainder are open-ended qualitative questions that are intended to enhance understanding of the quantitative questions.

Table 1
Theoretical Description of the Participant Feedback Questionnaire for Workshop 1

<i>Aspect</i>	<i>Definition</i>	<i>Number of questions</i>
Expectations regarding the workshop	Identify the concepts conveyed and estimate the degree to which they correspond to those desired by the participants	3 (2 open, 1 closed)
Perception of the educational value of the workshop	Ability of the workshop to transfer new concepts	2 (closed)
Overall satisfaction	Participants' assessment of the various program components	6 (closed)
Level of difficulty	Difficulty of workshop content	2 (closed and open)
Usefulness	Extent to which the concepts acquired during the workshop match the participants' professional needs	2 (closed and open)
Facilitator	Quality and skills of the facilitator	3 (closed and open)

These aspects correspond to those proposed by Kirkpatrick and Kirkpatrick (2006) for measuring participant reactions to training.

Psychometric Properties of the Questionnaire

Establishing the validity of a questionnaire involves applying procedures to demonstrate empirically that it measures what it is intended to measure. To this end, one needs to examine the degree of correspondence between its structure and participant responses.

An exploratory factor analysis was used for the purpose of this study. It consisted in conducting a principal component analysis to see if participant responses were consistent with the theoretical aspects of the questionnaire (e.g., expectations, educational value, satisfaction). Before conducting that kind of analysis, some statistical assumptions should be considered. The decision as to whether or not to use these statistical assumptions must be based on the aims of the validation study. In this case, it was to explore and gather evidence suggesting that participants' responses can be organized in accordance with the theoretical aspects covered in the questionnaire. The analyses were not designed to produce inferential knowledge to be used to predict participants' scores; this allows more tolerance in the application of statistical assumptions (Tabachnick & Fidell, 2007). The first assumption concerns sample size: it must be large enough to produce stable, reliable analyses (Tabachnick & Fidell, 2007). Consequently, we used data from the common questions from both versions of the feedback questionnaire (2003 and 2007). The number of respondents per question thus ranged between 87 and 197. Under these conditions, the analyses are expected to produce sufficiently reliable results. A second statistical assumption concerns the evaluation of outlier bias in the analysis. A few variables had extreme scores ($z > +$ or $- 3.29$). It should be noted that variance for each question was very low, increasing the likelihood of extreme scores. In this context, deleting participants with extreme values would have been equivalent to rejecting those who gave a different response. For these reasons, we decided to retain them for the analyses, despite the fact that doing so could make the results somewhat unstable.

The last assumption refers to the assessment of the *factorability* of the data, which means checking whether there is a probability that at least one factor can be extracted from them (Tabachnick & Fidell, 2007). A simple way to evaluate this is to observe a correlation matrix that includes all variables simultaneously. Visual examination of this

matrix indicates that the majority of them share a linear relationship with several others except for Question 3 (“The concepts presented by the workshop were new to me”), so it was withdrawn from the analyses. Its inclusion would have resulted in an unnecessary increase in statistical error (Tabachnick & Fidell, 2007). Furthermore, this procedure reveals that there is no singularity or multicollinearity. In other words, no pair of variables shows higher correlation than 0.75. This would indicate redundancy between such variables (no empirical distinction between them) and would unnecessarily increase the statistical error (Tabachnick & Fidell, 2007). Preliminary analyses have helped ensure that the data meet the statistical assumptions sufficiently for factor analysis to be used for the main analyses.

FINDINGS

SPSS 17.0 software was used to carry out a principal component analysis with varimax rotation on 11 questions with a sample of 197 respondents.³ Four factors were extracted using the principal component analysis,⁴ and each of these factors corresponded conceptually to the aspects covered by the questionnaire. As such, the solution decomposed into four factors⁵ that provided a very good representation of all the questions (variance explained by the solution = 85%) and demonstrated that each factor contributed to the solution.

The same approach was used to conduct a confirmatory factor analysis. It was based on participant responses to the questionnaire administered after Workshop 4. This questionnaire contained all the same elements as the one used for Workshop 1. Because the only items that differed were those regarding training content, we expected the confirmatory factor analysis to yield the same factors as the first analysis. To avoid repetition, only the highlights of the second factor analysis are presented.

A comparison of the first and second factor analyses reveals that the factors extracted were indeed identical. Discrepancies between the two analyses are minimal and can likely be attributed to the fact that most of the aspects in the questionnaire comprised only one or two questions, thereby increasing the instability of the analyses (Tabachnick & Fidell, 2007). However, this also strengthens the general validity of the questionnaire. Remember that only the type of knowledge transferred changed from one workshop to the next. Thus, the second factor analysis enabled us to determine that the organization of the participants' responses was not primarily influenced by

the nature of the knowledge transferred, but by other aspects of the training program. These findings suggest that the questionnaire is effective in measuring participant responses to all the relevant components of the training program.

DISCUSSION AND CONCLUSION

Over the years, responses by participants to the reaction questionnaires have resulted in adjustments to the training program and its improvement as a function of participant needs. For instance, when participants mentioned that there was too much theoretical content at the expense of practical exercises aimed at assimilation of learning, more time was devoted to the logic model. During the first training workshops, participants did exercises involving different programs, based on vignettes; following changes to the program, participants are now asked to divide into groups on the first day of the workshop and to choose a program that one of them is familiar with. This program becomes the subject of all exercises throughout their training, including the development of a logic model, an implementation evaluation plan, and a comparison group. As well, more time is now given to plenary discussions about the exercises, as participants indicated that they appreciated receiving feedback and hearing about the various challenges presented by their colleagues' programs.

In general, questionnaires on participants' reactions to training are designed to gather specific information to improve these activities. To do this, it is useful to assess the relevance of this type of instrument. This was the goal of our approach: to gather evidence that the questionnaire can provide valid information. For example, if one aspect is about the usefulness of the questionnaire, can we assume that the questions relating to this aspect represent a single factor or do they belong to another?

The validation study was used to document the psychometric properties of the participant feedback questionnaire administered in the context of the Essential Skills Series in Evaluation.⁶ In doing so, it delineated the questionnaire's relevance, structure and psychometric properties. This should facilitate the suitable use of the questionnaire and, eventually, help establish interpretive standards. In short, the validity assessment found that this instrument offers a satisfactory measure of participant reaction. Other activities are now required to assess the outcomes of the ESS training program in

terms of the knowledge acquired by participants and its impact on their evaluation practices. In describing our approach, we wanted to emphasize the importance and feasibility of assessing the validity of training program evaluation instruments.

NOTES

- 1 Because the data did not allow a test/retest procedure, the reliability of the questionnaire has not been evaluated.
- 2 Because the questionnaires for each of the workshops shared the same structure, only one is described.
- 3 Readers interested in a detailed example of procedures to be performed with SPSS to complete a factor analysis should consult the book *SPSS Survival Manual* by Julie Pallant (2001).
- 4 As indicated by the root sum square (RSS) of the factors, these showed good internal consistency and were well defined by the variables; the lowest RSS was 1.23. Conversely, the principal component extracted gave a good representation of each of the variables; its quality ranged from 0.73 to 0.92. With a criterion of $r = 0.45$ for accepting the inclusion of a question for the interpretation of the factors, all of these were retained. Further information regarding the statistical analyses is available upon request.
- 5 With one exception, all the questions reflect the aspects identified in the theoretical description.
- 6 The final version of the questionnaire for the first workshop day is included in the appendix. Versions for days 2 to 4 are similar.

REFERENCES

- Chiocchio, F. (1996a). A model for integrating theory and practice of psychometry. *International Journal of Psychology*, 31(3–4), 483.
- Chiocchio, F. (1996b). *Un modèle opératoire de création d'instruments de mesure*. Presented at the 9e Congrès de l'Association Internationale de Psychologie du Travail de Langue Française, Sherbrooke, Québec, Canada.

- Dutil, J., & Dagenais, C. (2008, May). *La formation de professionnels à l'évaluation de programme: bilan de l'expérience avec la Série des compétences essentielles de la SCÉ dans le cadre du symposium Enseigner l'évaluation de programmes: analyses réflexives*. Presented at the 2008 Conference of the Canadian Evaluation Society, Quebec City, Québec, Canada.
- Hogan, T. P. (2003). *Psychological testing. A practical introduction*. New York, NY: Wiley.
- Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating training programs: The four levels* (3rd ed.). San Francisco, CA: Berrett-Koehler.
- Pallant, J. (2001). *SPSS survival manual: A step by step guide to data analysis using SPSS for Windows*. Crows Nest, Australia: Allen & Unwin.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn & Bacon.

Christian Dagenais is associate professor of psychology at Université de Montréal. He was the Director of Evaluation at the Centre for Liaison on Intervention and Prevention in the Psychosocial Area (CLIPP) from 2002 to 2011. Over the last 15 years he has participated in the evaluation of numerous programs and has directed a theme segment of CJPE related specifically to the evaluation of knowledge transfer strategies. Dagenais is also co-author of a textbook on evaluation, *Approches et pratiques en évaluation de programme*.

Luc Dargis Damphousse is a doctoral student at the University du Québec à Montréal (UQAM) in the Department of Psychology. His research interests include methodological issues in quantitative and qualitative research, knowledge transfer, ageing, and program evaluation. He has an M.Sc. in psychology and experience working with a variety of populations in community-based research studies.

Julie Dutil holds a master's degree in organizational development and analysis. She is Director of Evaluation Projects at the Centre for Liaison on Intervention and Prevention in the Psychosocial Area (CLIPP), where she is responsible for all stages of evaluation of the various knowledge transfer projects of the centre.

Appendix
Final Version of Workshop Questionnaire, Day 1



Centre de liaison sur l'intervention
 et la prévention psychosociales

ESSENTIAL SKILLS SERIES
WORKSHOP EVALUATION QUESTIONNAIRE

Participant number: (Last four numbers of your personal phone number)
 Workshop date: ____/____/____
 Workshop location: _____
 Name of facilitator: _____

1. What was your primary objective when you enrolled in this training program?

For each of the following statements, circle the number that corresponds to your level of agreement.

1 = Strongly disagree	2 = Disagree	3 = Agree	4 = Strongly agree
------------------------------	---------------------	------------------	---------------------------

Training objectives and content

2. The content of this workshop met my expectations. 1 2 3 4
3. The concepts presented by the workshop were new to me. 1 2 3 4
- 3.1 If you circled 3 or 4:
 This workshop allowed me to fully integrate these new concepts. 1 2 3 4
4. The level of content difficulty was appropriate. 1 2 3 4
- 4.1 If you circled 1 or 2, specify: Too difficult Not difficult enough
5. The workshop provided me with information I can use in my work. 1 2 3 4

Workshop delivery

6. I am satisfied with the following elements of the workshop:
- 6.1 Course material 1 2 3 4
- 6.2 Audiovisual material 1 2 3 4
- 6.3 Food service 1 2 3 4

For each of the following statements, circle the number that corresponds with your level of agreement..

1 =Strongly disagree	2 = Disagree	3 = Agree	4 = Strongly agree
----------------------	--------------	-----------	--------------------

Teaching strategy

- 7. The exercises carried out contributed to my learning. 1 2 3 4
- 8. The number of exercises proposed was appropriate. 1 2 3 4
- 8.1 If you circled 1 or 2, you would have preferred: Fewer exercises
More exercises
- 9. Exchanges with the facilitator contributed to my learning. 1 2 3 4
- 10. The group discussions following the exercises contributed to my learning. 1 2 3 4

Facilitator

- 11. The facilitator was knowledgeable regarding workshop content. 1 2 3 4
- 12. The facilitator communicated clearly. 1 2 3 4
- 13. The facilitator gave satisfactory answers to questions. 1 2 3 4
- 14. The facilitator helped maintain my interest in the workshop. 1 2 3 4

Overall reaction

- 15. Overall, I was satisfied with this workshop. 1 2 3 4

16. What did you like most about this workshop?

17. What did you like least about this workshop?

18. Do you have any suggestions (content, material, other) to improve this workshop?

