

Evaluating an evaluation course

Professional training has long been the target for evaluation. However, most evaluation of training is confined to simple end-of-course data collection and superficial analysis. We argue that evaluations that rely heavily on end-of-course data collection *can* provide useful findings if they are well designed. This involves thinking clearly about the variables that are to be measured, translating them into measures, and ‘making sense’ of the findings within a framework set by the phenomenon under review. A course that presented training in evaluation to graduate students and participants in New Zealand was chosen as a case to exemplify the application of these design principles.

Introduction

Evaluation has had a relatively limited life as a disciplinary area in New Zealand and there has been an ongoing need for relevant knowledge and skills among those responsible for conducting evaluation work. From time to time, agencies in New Zealand offer courses that are part of related qualifications, or one-off workshops that take advantage of available expertise. The majority of people who carry out evaluations in public and private sectors or in academe have come from other disciplinary bases and used a mixture of practice and available courses to further their evaluation knowledge. Perrin et al. (2003) points out that ‘it can be at least as important for an evaluator to have interpersonal and communication skills as to have technical research skills’ (p. 236). However, there is a need to ensure that the highest standards of practice are fostered and maintained in a climate where the need for and number of evaluations of policy, practice and programs are increasing.

The course

Observers might expect that the routine review of evaluation training programs would be routine and that the literature would be redolent with examples of such evaluations. Unfortunately this is not the case as case studies of effective training would be useful for those about to develop training programs. This paper provides one such example.

A course titled ‘Program Evaluation: Concepts and Practice’ has been offered four times in Wellington during the past five years. Initially the course was offered as part of an AES Pilot Training Strategy. The Centre for Program Evaluation at the University of Melbourne presented this first course as the result of a competitive tendering process throughout Australasia. Subsequent courses were

Jenny Neale

Karilyn Andrew



Associate Professor Jenny Neale (top) is Head of School, Social and Cultural Studies, Victoria University of Wellington, Wellington. Email: <Jenny.Neale@vuw.ac.nz>

Karilyn Andrew (bottom) is an MA (Applied) in Social Science Research student in the School of Social and Cultural Studies, Victoria University of Wellington, Wellington. Email: <karilyn@paradise.net.nz>

presented in association with Victoria University of Wellington (VUW) and catered for a mix of students and those working professionally in the area of evaluation. The most recent course was held in April 2004. Our paper reports on the evaluation of this course and explores to what extent the aims of the course were met as well as discussing suggestions participants made for changes.

Participants

There were 24 participants on the course. Eleven were students in the Master of Arts (Applied) in Social Science Research programme at Victoria University of Wellington (the students) and the remaining 13 were practitioners (the practitioners), predominantly working in the public sector. This latter group entered the course via Victoria University's Centre for Continuing Education and Executive Development, Te Whare Pukenga (CEED).

Broadly, these two groups fitted the courses' intended target audience, that is, people who have some experience in evaluation, were intending to undertake evaluation work, or have a responsibility for evaluation studies and for those with experience in related areas such as policy development and analysis.

Study design

The evaluation of the evaluation course was primarily an 'impact evaluation' (Owen with Rogers 1999, p. 39) occurring at an end point after the course had taken place. Impact evaluations seek to assess the impacts of a particular program (Owen with Rogers 1999, p. 47), and are concerned with outcomes, which 'are benefits for participants during or after their involvement with a program. Outcomes relate to knowledge, skills, attitudes, values, behavior, condition or status' (Owen with Rogers 1999, p. 264).

Towards the end of the last day of the five-day course, two evaluation instruments, were administered, one on behalf of CEED and the other from the course presenters. A total of 47 questionnaires were completed as one practitioner only filled out the questionnaire from CEED. Participants were given the option of making their submissions anonymous, but were asked to indicate whether they were students or non-students. They were also told that one of the students who was also a course participant would prepare a report on the findings as an assignment for her higher degree. With this knowledge, participants then made their choice as to whether they put their names on the two questionnaires; 18 participants did so.

After the questionnaires were completed, the course presenters facilitated a discussion where

participants had an opportunity to give verbal feedback about the course. The graduate student took extensive notes during this session, which lasted for 50 minutes. The results presented in this paper come from both sets of questionnaires, and the feedback session.

Evaluation findings

Knowledge and understanding of evaluation principles

The emphasis of the questionnaires was on knowledge acquisition of participants and their expected use of the knowledge. The advantages of this type of evaluation are that it can be ascertained 'how relevant participants thought the training was ... whether they were confused by any of the training ... point out any areas in which trainees thought information was missing ... tell us how favorable overall participant reactions were' (Krein & Weldon 1998, p. 17).

Participants were asked to rate their knowledge and understanding of evaluation principles on a scale from 1 (low) to 10 (high), both before and after the course. This gave an indicative self-assessment of the effect of the course.

As can be seen in Figure 1, the range of practitioners' responses about their initial knowledge was from 2 to 8 compared to the Student range of 0 to 6. Both groups indicate a wide range of knowledge and understanding. The Practitioners range reflects their diverse experience, from that of senior evaluators, to those just becoming involved in evaluation practice. The Student responses also reflect a range of experience but start from a lower base.

Figure 2 shows the means for the Student group and the Practitioner group. As could be expected, the higher mean for the Practitioners indicates some working knowledge and understanding of evaluation principles acquired prior to the course. Students, on the other hand, are less likely to have been involved to the same extent or in the same capacity in the field of evaluation.

Figure 3 illustrates ratings that participants gave about their knowledge and understanding at the end of the course. The range of Practitioner responses was from 5 to 8.5, with the Student range being from 5 to 8.

Figure 4 indicates that the gap between the two groups in terms of their knowledge and understanding of the principles of evaluation had closed considerably as a result of the course. Moreover, as the two groups had started at different levels, it could be said that Students had more new knowledge and understanding to assimilate than the Practitioners, hence they gained more. However, both groups of participants perceived that their knowledge and understanding of evaluation

FIGURE 1: RATING OF KNOWLEDGE AND UNDERSTANDING AT THE BEGINNING OF THE COURSE

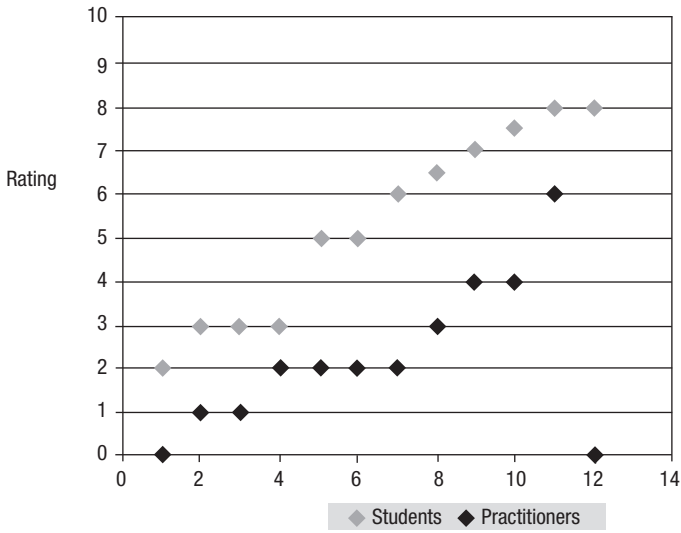


FIGURE 2: MEAN VALUE OF KNOWLEDGE AND UNDERSTANDING AT THE BEGINNING OF THE COURSE

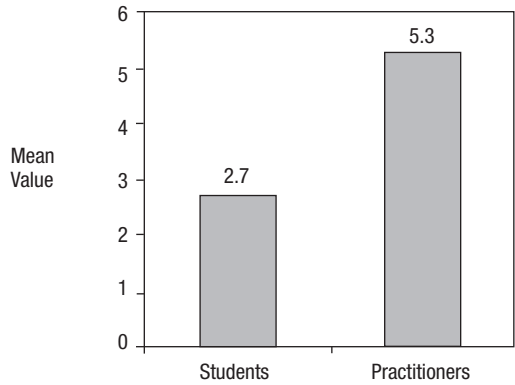


FIGURE 3: RATING OF KNOWLEDGE AND UNDERSTANDING AT THE END OF THE COURSE

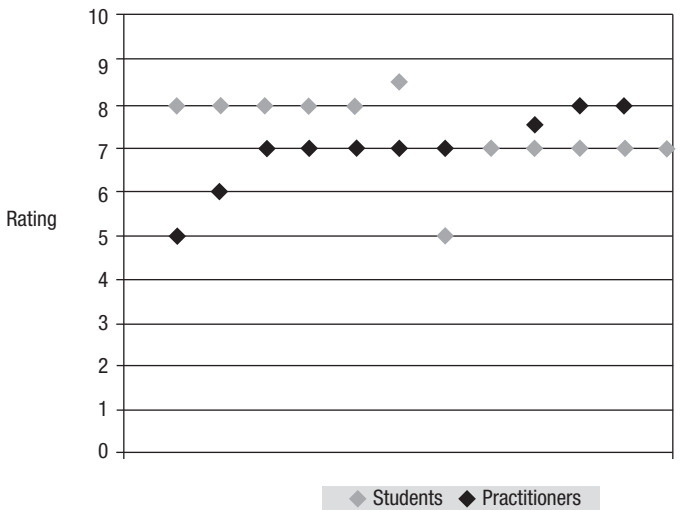


FIGURE 4: MEAN VALUE OF KNOWLEDGE AND UNDERSTANDING AT THE END OF THE COURSE

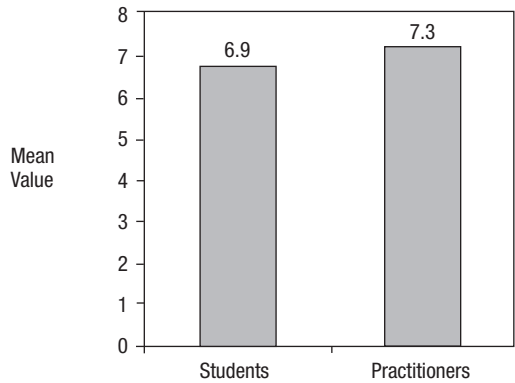


FIGURE 5: RATING OF ABILITY TO COMMISSION AND/OR CARRY OUT AN EVALUATION AT THE BEGINNING OF THE COURSE

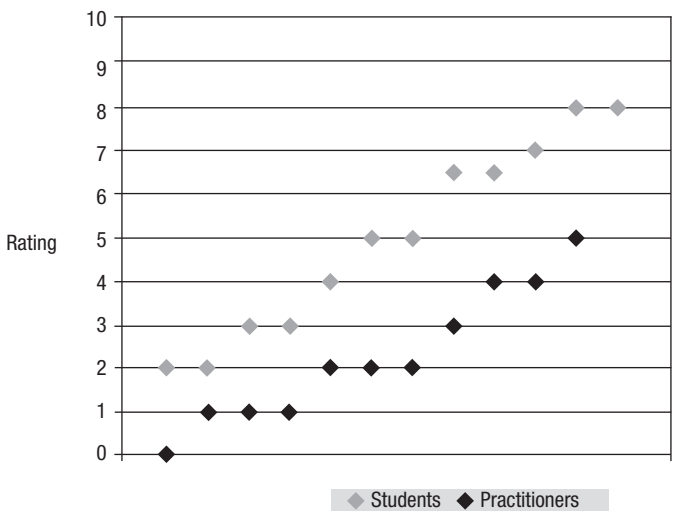


FIGURE 6: MEAN VALUE OF ABILITY TO COMMISSION AND/OR CARRY OUT AN EVALUATION AT THE BEGINNING OF THE COURSE

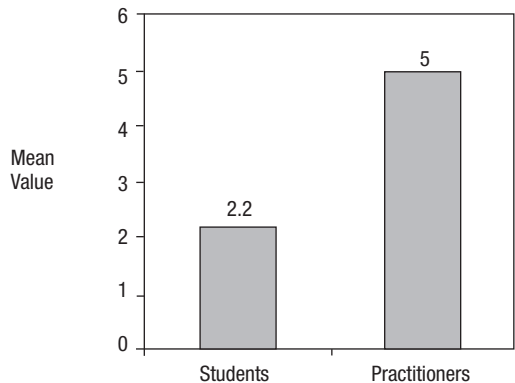


FIGURE 7: RATING OF ABILITY TO COMMISSION AND/OR CARRY OUT AN EVALUATION AT THE END OF THE COURSE

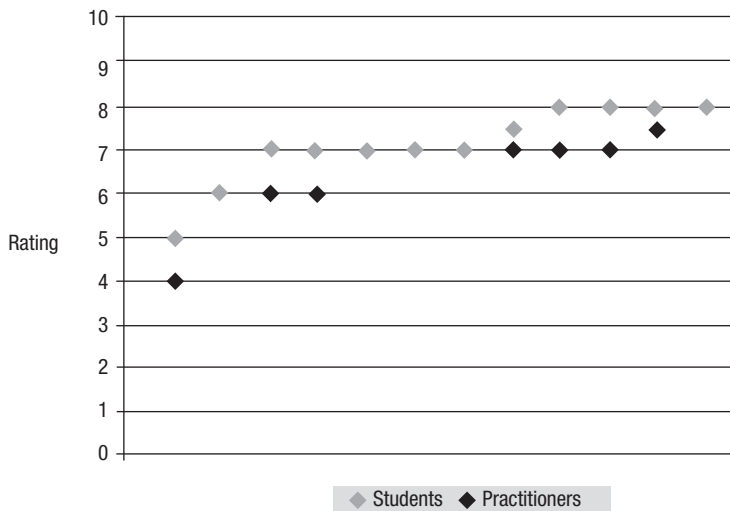
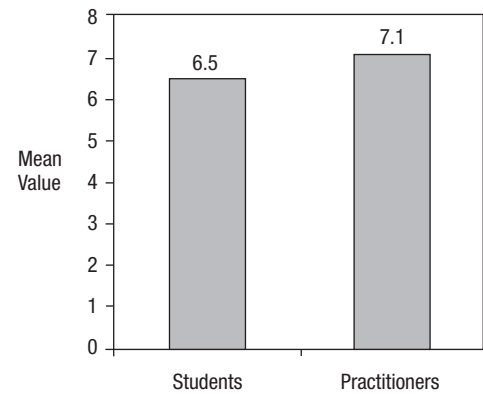


FIGURE 8: MEAN VALUE OF ABILITY TO COMMISSION AND/OR CARRY OUT AN EVALUATION AT THE END OF THE COURSE



principles increased markedly as a result of participating in the course.

Ability to commission and/or carry out an evaluation

At the end of the course, participants rated what they perceived their ability was at the beginning of the course, to commission and/or carry out an evaluation. The Student range depicted in Figure 5 is from 0 to 5, with the Practitioner range being much wider and higher from 2 to 8. This difference is comparable with the difference between Practitioners and Students and their understanding of evaluation principles at the beginning of the course.

Figure 7 shows that after participating in the course the gap between Student and Practitioner ratings of their ability to commission and/or carry out an evaluation has closed. Figure 8 illustrates the mean values, showing that the Practitioner group rated themselves only marginally higher overall than Students did. This is comparable to the situation of the mean values regarding knowledge and understanding of evaluation principles after participation in the course.

Students perceived they had less ability to commission and/or carry out an evaluation prior to the course than Practitioners did and again appear to have gained more from the course because they began at a lower level. However, both groups of participants perceived that their ability to commission and/or carry out an evaluation had increased as a result of participating in the course.

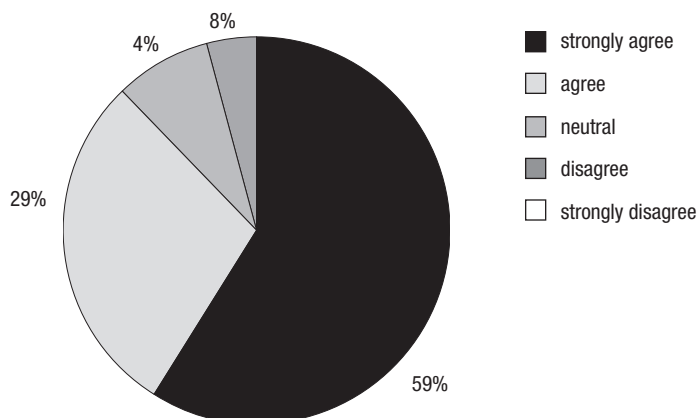
It could be tempting to interpret these results in a way that would suggest that the course was able to

redress inequalities of knowledge and understanding of evaluation principles, and abilities to commission and/or carry out evaluations between Students and Practitioners. Kirkpatrick (1998) would argue that this, however, would be an incorrect deduction, because no actual measurement of learning has taken place, but rather only the measurement of perception (1998, p. 4). Phillips, however, makes a case for self-assessment by participants of their learning, arguing that ‘in many applications, a self-assessment may be appropriate, by which participants are provided an opportunity to assess the extent of skills and knowledge acquisition’ (1997, p. 131). While it is possible that participants are good judges of their own learning, it is still not evidence that Practitioners and Students per se, have the same knowledge and understanding in regard to evaluation principles, nor the same abilities in commissioning and/or carrying out evaluations.

Furthermore, these results do not measure actual changes in participants, in terms of what is demonstrated by on-the-job behaviour. For this, a strategy within the evaluation (Kirkpatrick 1998, p. 5) that essentially attempts to measure the ‘transfer of training’ (Garavaglia 1998, pp. 74–77) from the course to the actual workplace or practice (Kirkpatrick 1998, p. 5) is required. However, such evaluations are problematic because of the difficulty of determining cause and effect when trying to isolate the influence of such courses from other factors contributing to change (Kirkpatrick 1998, p. 7; Kirkpatrick 2004, pp.10–11, Queeney 1996, p. 719).

In spite of the above cautions, as a result of participating in the course, participants perceive that

FIGURE 9: RESPONSES TO THE STATEMENT: 'THE TEACHING WAS EFFECTIVE'



their knowledge, understanding and abilities have increased with regard to evaluation principles, and in commissioning and/or carrying out an evaluation.

Findings about processes

A comprehensive impact evaluation not only examines outcomes for participants but also collects evidence about the processes that might explain the pattern of outcomes. Within the restrictions placed on the study by resources, some process or implementation data were collected and analysed. The first of these was participant opinions about the teaching.

Effectiveness of teaching

Overall, 88% (21) of participants either Strongly Agreed or Agreed with the statement that the teaching was effective. See Figure 9.

Group work

Group work was used extensively during this training program on exercises that were drawn from a case study of a simulated evaluation. There is evidence (Harbour 1998) that group work has been rated as one of the most favoured instructional methods. In this example, more Students than Practitioners commented on group work, and for the most part, the Students were very positive about it, both in terms of the exercises themselves, and the benefits of working in groups. One made the point that ‘group activities involving all members in producing work in a given time frame is an excellent way to learn’. Other Students indicated that the mix of participants was important in the group work. For example, ‘I felt that the group exercises were very beneficial as I learnt a great deal from talking to those involved in the evaluation industry’. These comments suggest that using this teaching approach

was clearly useful for Students, as it is based on adult learning principles, where, as Knowles indicates ‘adults bring into a learning situation a background of experience that is itself a rich resource for many kinds of learning for themselves and for others’ (1996, p. 256). A Practitioner agreed, stating that they ‘found the exercise working in the groups most useful because group members had prior knowledge of evaluation’ which suggests that Students were not the only ones who thought they benefited from others’ experience.

However, Students felt that the mixed groups were not always useful. For example, one Student commented that ‘there was a tendency within the two groups that I was in for the “workers” to privilege their own knowledge over the students which was actually unfounded at times. Students had life experiences that were a valid contribution’. Knowles explains that ‘if adults’ experience is not respected and valued, is not made use of as a resource for learning, they experience this omission not as a rejection of their experience but as a rejection of themselves as persons’ (1996, p. 256). Further to this, at least one Practitioner expressed that the ‘mix of students and experienced evaluators did not work well, the course was pitched too low for those with some experience with the result that some sessions were not very productive or dragged’. They thought that ‘more teaching would have been better’. This suggests that the diversity of experience levels within the group context was problematic for some participants, and not necessarily a positive experience at times.

Other Practitioners alluded to tensions in group dynamics, but they did not specifically attribute them to the Student/Practitioner mix. ‘I really enjoyed it, though some of the group discussions got a bit too intense at times.’ Another implied, albeit indirectly, that not all group work was necessarily ‘smooth sailing’: ‘there was a large proportion of class time devoted to group work, we were lucky that our group got on well and was productive’.

Some of the students commented that although group work may have been enjoyable or useful, at times there were preferences for a greater mix of teaching methods; for example, ‘excellent practical exercises for groups to engage in. I would have liked some more individual exercises at times’, and ‘enjoyed the group work, although sometimes working in pairs was more efficient’. One Student, however, ‘would have liked more structured class time. I personally don’t like/respond well to interactive group situations. Often one or two people will dominate’.

The group work essentially involved participants in practical exercises, and was predicated on adult learning principles, enabling participants to ‘tap into the accumulated knowledge and skills of the

learners' (Knowles 1996, p. 256). Overall, more Students than Practitioners commented positively about the group exercises and the benefits gained, though the majority of Practitioner comments, like those of Students, were positive.

In discussing the use of interactive learning techniques Thiagarajan notes:

These instructional techniques are based on two important premises: (1) People learn better through active experiences than passive listening; and (2) people learn better through interacting with one another than working alone. (1996, p. 517)

Knowles (1996), Thiagarajan (1996), and Harbour (1998) all emphasise people's active participation in their own learning. The use of groups to achieve this in the context of the course was not without its tensions. These mainly related to group dynamics, with the intensity of discussions/arguments noted by one participant, and others noting the potential for some members to dominate as well as the invalidating of Student experience and knowledge.

Theory and practice

A major component of the teaching approach of the course was its practical nature, and orientation towards issues that are of relevance to the workplace with conceptual frameworks helping to make more sense of existing practice.

The participants who commented on this teaching approach using both theory and practice in tandem were predominantly Practitioners and were universally positive. Students tended to comment about either practice or theory as separate components. Practitioners thought that it was a 'good mix of practical and theory', and a 'good, quick snapshot of evaluation theory and practice'. One thought that 'it was hands-on, with theory supplementing what was learned during workshop sessions' and that it 'provided a very good outline of evaluation concepts and the issues arising from evaluations'. While another considered there were 'extremely well-designed practical exercises—the best I've ever seen in a course'.

A Student also commented on the exercises, noting the 'strong focus (through the activities) on applying knowledge'. They thought it was a good approach for the course and appropriate for participants because they perceived that in the New Zealand context there were 'quite a lot of evaluation projects being commissioned and few experienced evaluators'. They also thought the Course Notes were well linked with the activities, which illustrates Phillips' contention that handout materials 'can support the information presented in the program and provide for additional analysis

and follow-up when the participant is back on the job' (1997, p. 293).

Participants described a learning environment created by the instructors where the course was 'very well presented in a relaxed, pleasant manner', with a 'great teaching style. I was focused and attentive most times'. Another commented that 'it was lively and interesting. The tutors were easy with the class and it was most enjoyable'. The main presenter was described as 'a good facilitator', 'an excellent tutor/presenter'. The team approach to teaching was noted: 'Jenny and John worked very well as a team', and that it was 'good having both teachers as it kept up the interest'.

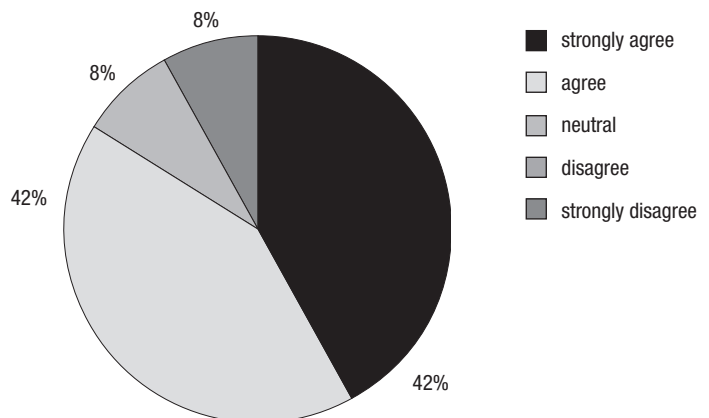
Students and Practitioners commented equally positively about the instructors' knowledge and expertise. A Student noted that because the teachers 'were well experienced in their field of expertise' it made 'it more of a transfer of knowledge than if it had all been theory based'. Participants found the real-life examples very valuable. For example, 'the practical real-life examples of the peaks and troughs of evaluation at all stages were especially enlightening'. 'Analogies also help increase training transfer, by showing how important principles can apply in various situations' (Garavaglia 1998, p. 76).

Comments about the quality of the course, and the teachers reinforce Harbour's findings that 'the ability to create a learning environment was judged the most important characteristic, with competency in the subject matter coming in a close second' (1998, p. 119).

Content pitch

Overall 84% of the course participants indicated that they either Strongly Agreed or Agreed with the statement that the content was pitched at the right level. See Figure 10.

FIGURE 10: RESPONSES TO THE STATEMENT: 'THE CONTENT WAS PITCHED AT THE RIGHT LEVEL'



Some of the Students were unsure whether it was pitched well for both groups, as one said it was right ‘... for me being new in the field. For other experienced practitioners it may have been a bit basic’. However, responses such as ‘found it very useful, especially from a practitioner point of view’ and ‘catered to the diversity of the participants—both in terms of their knowledge and their experience’ suggest that for some Practitioners the course was indeed pitched at the right level.

One student commented, ‘because the class had varying levels of expertise—some were more out of their depth’. A Practitioner responded in a similar vein, noting that the ‘variety of participants’ experiences means not everyone will agree’ in regard to whether content was pitched at the right level.

Generally, the feedback suggests that the course was pitched well for the majority of the participants, both Practitioners and Students. Comments from both questionnaires showed that on either end of the continuum, there were issues regarding the pitch of the content. Students with minimal knowledge of the subject prior to the course felt that it was pitched too high, and very experienced Practitioners, with a number of years experience as evaluators, felt it was pitched too low.

Content

Figure 11 shows that 86% of all participants on the course Strongly Agreed or Agreed with the statement that the course content was well organised.

The majority of Practitioner responses related to professional development needs

that they perceived they had, and have for the future. The quotes illustrate that ‘ideally this education enables practitioners to keep abreast of new knowledge, [and] maintain and enhance their competence’ (Queeney 1996, p. 699). For example, one Practitioner stated that ‘the course was an excellent coverage of the field, and allowed me to refresh my knowledge of where evaluation had developed and where the basic principles had remained fairly stable ... Confirmed for me what I know and areas where I needed to improve or develop new skills such as the use of software programmes for quantitative analysis’.

A Student also commented that they were also now more comfortable in terms of carrying out a small-scale evaluation, and that they were ‘really impressed with the amount of content squeezed into this week’. A fellow Student felt that the ‘application of content has been steady and at a good pace. My interest in evaluation has increased’.

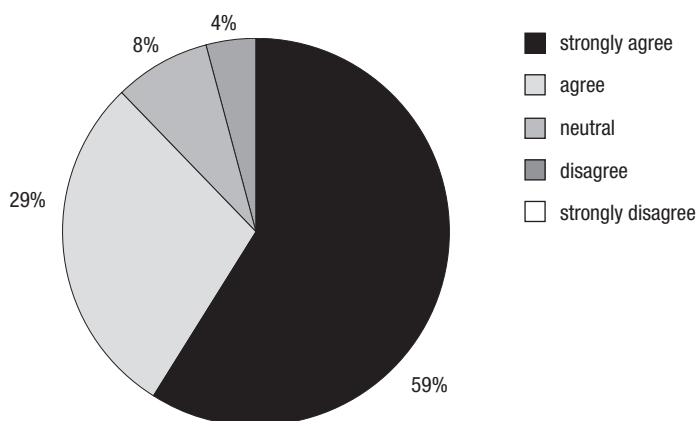
Participants rated the course highly both in terms of overall quality and in giving the course an overall assessment. The mean participant rating of the overall quality was 8 on a 10-point scale, and 79.5% of participants rated the course overall as either Excellent or Very Good.

Key findings

Overall, participants perceived that their knowledge and understanding of evaluation principles, and their ability to commission and/or carry out an evaluation increased as a result of participating in this course. Participants particularly liked the expertise and real-life examples and anecdotes that the course instructors used, as well as their general manner and open approach. This affirmed the approach taken in delivering the course to achieve its objectives, which included it being ‘oriented towards practical issues which are of relevance to real and current problems found in the workplace’ (Owen & Neale 2004).

Being based on adult learning principles enabled the use of a case study ‘to introduce and reflect on key evaluation concepts and frameworks’ along with the ‘emphasis on group work throughout the course’ (Owen & Neale 2004). Thus group work was a major component of the teaching approach, and Students who commented were particularly focused on its benefits in terms of gaining from the experience which others, particularly the Practitioners, brought to the groups. Practitioners signalled that this was not a reciprocal feature of the group work, and overall they were less positive about the Student/Practitioner mix of the groups.

FIGURE 11: RESPONSES TO THE STATEMENT: ‘THE COURSE CONTENT WAS WELL ORGANISED’



There were tensions in some of the groups, not directly ascribed to the Practitioner/Student mix although this was not precluded as a cause. As a way to ease tension, making changes in the composition of groups was suggested. Some wanted the opportunity to interact with more people, whereas others wanted to work with different people. For example, one participant suggested ‘perhaps the groups could have been rotated more often and size changed—it was in some cases easier to work on some of the activities in pairs or smaller groups.’

Issues relating to the pitch of the content may have contributed to Practitioners being less positive than Students about the group work. While 91% of Students Agreed or Strongly Agreed that the content was pitched at the right level, only 77% of Practitioners did so.

The majority of participants, however, indicated that the content was pitched well and the teaching approach of theory and practice through the conceptual frameworks and practical exercises in groups was appreciated. This met the objective of ‘conceptual frameworks [being] provided that will help make more sense of existing practice’ (Owen & Neale 2004).

Practitioners commented about the theory and practice approach combined, whereas while Students generally did not make explicit links between the practical and theoretical components, Practitioners tended to do so. Participants agreed that they found value in the practical exercises and having an overview of the theory and concepts at the beginning of the course. Changing the structure of the course so that the theory was presented first, rather than simultaneously with practice, was suggested as a way of diffusing perceived tension. More depth or detail in the theory and practice aspects of the course was also advocated.

Other specific examples of content participants would have liked in the course included: more discussion of the New Zealand context; some information for students on how evaluation operates in ‘real-life work environments’; an exercise in setting out the questions of a scenario evaluation, for example role-playing the negotiation phase; an exercise in commissioning evaluations and steps needed to ensure that the contracted evaluators stay focused and on the right track; and how to write evaluation reports. These suggestions, of course, besides considerably lengthening the course, indicate possible content for subsequent courses and highlight the need for ongoing professional development and training in the New Zealand context.

Conclusions

The evaluation of this evaluation course, besides providing an analysis of its content and form, and suggestions for further professional development and training initiatives, provides an interesting example of group dynamics. The course participants were from two different but not necessarily disparate groups, though for the purposes of this evaluation they were labelled so as to distinguish them as Students and Practitioners. The convenience of identifying them this way emphasised difference at a group level and thus minimised the evident diversity within the two groups and similarity across the two groups. For example, the majority of the Students were of a similar age range to the Practitioners and there were overlaps in terms of experience. Using adult learning principles to ensure that individuals are taken from where they are and can build on their knowledge base implies that the process is one of partnership in the teaching–learning endeavour. On the whole, participants on the course appeared to subscribe to this ethos. However, where this was not the case, participants tended to behave as if they ‘knew it all’ and could learn little from other participants (particularly if they were students). Ensuring more structured roles within group activities and having homogenous groups for some tasks may be worth exploring for future course delivery and, as a more general principle, when course participants are diverse.

Note

The original analysis for this evaluation was undertaken by Karilyn Andrews and the primary audience for the findings are Drs Neale and Owen who presented the course that is the subject of this evaluation study.

References

- Garavaglia, PL 1998, ‘How to ensure transfer of training’, in DL Kirkpatrick, *Another look at evaluating training programs*, American Society for Training & Development, Alexandria, Virginia.
- Harbour, JL 1998, ‘Does the trainee know best’, in DL Kirkpatrick, *Another look at evaluating training programs*, American Society for Training & Development, Alexandria, Virginia.
- Kirkpatrick, DL 1998, ‘Great ideas revisited’, in DL Kirkpatrick, *Another look at evaluating training programs*, American Society for Training & Development, Alexandria, Virginia.
- Kirkpatrick, DL 2004, ‘Evaluating training programs: evidence vs. proof’, in ‘Owen course notes: program evaluation—concepts and practice’, unpublished.

REFEREED ARTICLE

- Knowles, MS 1996, 'Adult learning', in RL Craig (ed.), *The ASTD training and development handbook: a guide to human resource development*, McGraw-Hill, New York.
- Krein, TJ & Weldon, KC 1998, 'Making a play for training evaluation', in DL Kirkpatrick, *Another look at evaluating training programs*, American Society for Training & Development, Alexandria, Virginia.
- Owen, JM & Neale, J 2004, 'Course notes: program evaluation—concepts and practice', unpublished.
- Owen, JM with Rogers, PJ 1999, *Program evaluation forms and approaches*, 2nd edn, Allen & Unwin, St Leonards, NSW.
- Perrin, B et al. 2003, 'Some thoughts and reflections from Europe, Australia and the UK' in N Lunt, C Davidson & K McKegg (eds), *Evaluation policy and practice: a New Zealand reader*, Pearson Education, Auckland.
- Phillips, JJ 1997, *Handbook of training evaluation and measurement methods*, 3rd edn, Gulf Publishing Company, Houston, Texas.
- Queeney, DS 1996, 'Continuing professional education', in RL Craig (ed.), *The ASTD training and development handbook: a guide to human resource development*, McGraw-Hill, New York.
- Thiagarajan, S 1996, 'Instructional games, simulations, and role-plays', in RL Craig (ed.), *The ASTD training and development handbook: a guide to human resource development*, McGraw-Hill, New York.