Evaluating course evaluations: the need to establish what is being measured

This item was submitted to Loughborough University's Institutional Repository by the/an author.


Additional Information:

- This article was published in the journal, Assessment and Evaluation in Higher Education [© Taylor and Francis] and the definitive version is available at: www.tandf.co.uk/journals/titles/02602938.asp.

Metadata Record: https://dspace.lboro.ac.uk/2134/2991

Publisher: © Taylor & Francis

Please cite the published version.
This item was submitted to Loughborough’s Institutional Repository (https://dspace.lboro.ac.uk/) by the author and is made available under the following Creative Commons Licence conditions.

For the full text of this licence, please go to:
http://creativecommons.org/licenses/by-nc-nd/2.5/
EVALUATING COURSE EVALUATIONS: THE NEED TO ESTABLISH WHAT IS BEING MEASURED

Jenny A. Darby
Department of Social Sciences, Loughborough University, Loughborough, Leicestershire. LE11 3TU
E mail J.A.Darby@lboro.ac.uk

The study attempted to assess the validity of measures commonly used for course evaluations. Activity and attitudinal measures were included and these were related to two different control groups. The study was completed with a sample of teachers who attended a course dealing with child abuse repeated 23 times with groups of approximately 25 teachers. It was found activities undertaken after the course were more a reflection on how little effort was required to carry them out than on the effectiveness of the course. Attitude measures tended to make it appear the course had an effect. The type of control group used for comparison purposes had a marked effect on whether the course appeared to have been effective. The need to consider these factors as influential variables when interpreting data from course evaluations was stressed.

Introduction

It was the aim of this study to assess the validity of some measures commonly used to evaluate training courses. These include behaviour changes and changes in attitudes or beliefs after a course. These measures were examined in the context of comparison or control groups.

Activity measures

The work of Browne & Meuti (1999) stress the success of an instructional workshop lies in its impact on the student’s behaviour some time after the workshop, rather than the immediate reaction of the participants. They suggest this can be assessed in terms of the activities the students have engaged in since the course. Schwarz et al. (1985) explain that the typical method used is to list response alternatives taken from a total list of activities promoted during the course. The assumption here is the more the participant does after the course the better the course. This assumption does not, in the present author’s view take into account motivational theory. Vroom (1964) looked at motivation with workers and found the amount of effort they expended was related to the return they got. This idea has been refined and developed by many researchers including Siegrist (1996). He has shown how workers try to maintain the balance between effort and reward. Siegrist suggests that if the balance is not maintained stress levels increase. On the basis of Siegrist’s (1996) effort-reward principle it would be expected if an activity recommended in a course requires a great effort it is less likely to be carried out. Effort could be intrinsic to the task or could be influenced by many other factors such as the need to co-operate with others. The idea of activity measures as a means of evaluation appears over-simplistic.

Attitude measures

Smith & Beno (1993) in their model of evaluation referred to attitude change by the participants as being an important long term effect of a course because of its link with behavioural change. Whether this assumption can be justified from the research evidence is debatable. Eagly & Chaiken (1993) describe an attitude as 'an inferred state with observable cognitive, affective and behavioural responses'. Wicker (1969) reviewed a number of studies and concluded attitude change did not predict subsequent behaviour. Kahle & Berman (1979), on the other hand, showed that an attitude can predict specific behaviour change if the occurrences are aggregated over a period of time. In addition, more recently, Kraus (1995) and Gibbons et al. (1998) showed that if a number of confounding influences can be neutralised it is possible to use a knowledge of attitudes to make accurate predictions of behaviour. In spite of this disagreement, many organisations use attitude change as a means of assessing the effects of training courses. A typical example is a study by Tait & Purdie (2000) in a topic area closely related to the one to be used in the present study. They found that the impact of a teacher training programme which included material on children with disabilities produced minimal attitude change. It is not intended in this study to become involved in the attitude/behaviour debate, but as attitudes are commonly used as a measure of course effectiveness, they are used as an evaluation measure in this study.
Control groups
Comparing the evaluation reports of different courses is common practice. Rae (1997 p. 77) provides a clear idea how control groups should be used in training programmes for this purpose. He explains that the group which receives the training should be tested before and after the course. The control group should be tested at the same times, but should not receive the course. In the ‘real world’ of training this procedure cannot always be followed. Access to participants before courses is rare. The exercise would in many instances have little value if the participants were to learn something about which they actually knew very little beforehand. The difference between no knowledge and even a small amount would be significant. Further this procedure would mean evaluation forms could not be anonymous. In practice course sponsors tend to use other types of control group. One consists of employees doing the job concerned but with no formal training. This meets Foster & Parker’s (1999 p. 26) requirement that it is ideal to have a control group formed from the same group of people as took part in the experimental group. Another compares the course participants against a work group the members of which had neither received training, nor were they doing the job. This meets Hayes’s (2000 p. 41) view that it is acceptable simply to match the participants in the two conditions. The use of these two types of control group has some academic support. Breakwell et al. (2000 p. 77/8) for example, argued that when it is not practicable to conduct a proper experiment with a control group it is acceptable to use some sort of ‘comparison group’.

In the present study a training course was designed and presented which was intended to raise awareness amongst child abuse co-ordinators in schools and help facilitate a requirement for all schools to create and implement a policy for dealing with child abuse. The course was approved and sponsored by the Local Education Authority and was mandatory for all Child Abuse co-ordinators located within every school in the geographic county. The course was repeated 23 times with class sizes of about 25 and consisted of three one day sessions held within a period of a few weeks. The training programme was designed to follow on from previous inputs providing an awareness of signs and symptoms. The training programme involved tutor presentations, groupwork and individual activities. In the course participants were told of, and provided with, a whole range of activities and procedures which they could use to cope with child abuse issues. Some of these had the aim of passing on information to ‘significant others’ about how to deal with child abuse. These ‘others’ included the children in school, their colleague teachers and outside agencies.

For the purposes of the present study an evaluation questionnaire was distributed to course participants three years after the course had been completed. Activity and attitude measures were incorporated into the evaluation and two control groups were included in the study.

It was hypothesised that:
1. The course participants would be more likely to use teaching materials which required less involvement/effort on their part than material which required more involvement/effort. (following motivation theory and Siegrist 1996).

2. The course participants would be more likely to carry out activities which did not require the co-operation of others than activities which did require the co-operation of others. (following motivation theory and Siegrist 1996).

3. The course participants would display more positive attitudes towards their own role and school procedures for dealing with cases of child abuse than members of two control groups who had not attended the course.

METHOD
Timing of the study
The evaluation of the course was carried out three years after it had been held for a number of reasons. The first was to allow time for In Service Educational and Training Sessions (INSETS) to be organised and run. In many schools they are planned twelve or more months in advance. The second, was to take into account the need, according to Petty (1994), to allow time for peripheral attitude change to take place. Third, due to staff turnover, a longer time period could have resulted in too few course participants still being in post.

Participants
Questionnaires were sent to the 529 teachers who were named child abuse co-ordinators in an East Midlands county. These consisted mainly of head and deputy head teachers. 273 were returned from those
who had attended the child abuse course. A further 101 were returned from co-ordinators who had not been on the course. These formed the first control group, meeting Foster & Parker’s (1999) requirement. Questionnaires were distributed and returned to the researcher by means of the schools internal mail system. Of the 527 questionnaires distributed 381 were returned. The response rate was therefore 71%.

A second control group consisting of qualified teachers, contacted during two teacher conferences was included to meet Hayes’s (2000) requirement. Of those given questionnaires 360 who satisfied the following criteria were included in the study. They had been teaching for at least three years (mean 3.7 years) and were not child abuse co-ordinators, nor were they from the schools included in the main sample. They had some basic knowledge of child abuse procedures from their Post Graduate Certificate in Education (PGCE) basic teacher training course.

THE EVALUATION QUESTIONNAIRE
Preparing the questionnaire
This was a self administered mail survey. A pilot study was undertaken using ten teachers who had been on the course. They all had experience of dealing with child abuse cases, worked in a number of different schools, and were not be involved in the final sample. As a result of the pilot study only minor alterations were required to the questionnaire which is outlined in the following section presented here, for clarity, linking each measure with the related hypothesis.

Hypothesis 1 The course participants would be more likely to use teaching materials which required less involvement/effort on their part than material which required more involvement/effort.

Teaching materials measures
This lists teaching materials promoted in the course (see table 1 and the appendix) which were available for use by schools.

Scale for effort required to use teaching materials
Following the procedure of Holmes & Rahe (1967) for creating a scale, the ten teachers in the pilot sample were individually allowed to familiarise themselves with the materials, to be used as a measure in the evaluation. They were given cards and asked to rank the materials in order, with that involving the least teacher effort on the top, then the second amount of effort and so on. The scale produced is shown in table 1.

<table>
<thead>
<tr>
<th>Title of materials</th>
<th>Mean rank order.</th>
<th>Lowest score is easiest for teacher to administer.</th>
<th>Standard deviations in brackets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kidscape</td>
<td>1.4 (0.52)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stranger danger</td>
<td>1.6 (0.52)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teenscape</td>
<td>3.2 (0.42)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Where to go for help</td>
<td>4.4 (0.52)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assertiveness training</td>
<td>5.0 (0.67)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>How to deal with harassment</td>
<td>5.6 (0.84)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Hypothesis 2. The course participants would be more likely to carry out activities which did not require the co-operation of others than activities which did require the co-operation of others.

There follows the measures included in the questionnaire which are referred to in relation to this hypothesis. These are presented in more detail in the appendix.

Activities carried out under the individual’s control
This enquired about activities which had taken place in the school and included revision of school guidelines, use of material for child protection training and attendance by staff on any training courses concerning abuse, other than INSETS. The data used was based on the total numbers recorded by each participant.
Activities not under the individual’s control
This enquired about the number of referrals made for children who may have been abused and refers to the total in the school not the number made by the co-ordinator. The number of referrals could be influenced by the agreement of other agencies such as the Social Services.

This section also enquired about the implementation of INSETS another measure not under the total control of the co-ordinators. Whether these took place could well, for example, be influenced by the school’s INSET programme and policy.

Additional questions asked co-ordinators in schools where INSETS had not taken place to rank in order of importance their view of the reasons why this was so.

Hypothesis 3. The course participants would display more positive attitudes towards their own role and school procedures for dealing with cases of child abuse than members of two control groups who had not attended the course.

Included in the questionnaire are statements which concern a range of feelings and attitudes towards child abuse and protection. The statements were designed to take into account the need to be suitable for both the control groups, who had not attended the courses. All responses were on a four point Likert scale from strongly agree to strongly disagree and were graded from one to four. Four indicated the most positive response. The four point scale of agreement was used in order to prevent avoidance responses by the use of a neutral option. Wells and Marwell (1976 p 83) argue that having a neutral point in a response scale provides a “golden mean” or the opportunity for an evasive response style particularly if the topic is an emotive or sensitive one. Some statements were reversed to avoid ‘response set’. The variation of means in tables 5 and 6 between categories reflects the greater number of questions in some categories. The four categories and the rationale for the questions are details below:

Feelings about knowledge of what to do
Question 1. I know what is required of me when dealing with a child who has or may have been abused.
Rationale - It was intended in the course that teachers should be given a good grounding of the procedures, and be made aware of what they should and should not do when faced with a child who may have been abused.

Acceptance abuse takes place
Question 2. A teacher’s job is to teach, not to act as a social worker.
Question 3. Child abuse referrals are given too high a priority in my school.
Question 4. I have a legal and professional obligation to refer suspected abuse to my co-ordinator
Question 5. If a child is not presenting any problems, suspected abuse is better ignored
Rationale - Lawlor (1993) found denial of abuse was common amongst teachers and so these four statements were related to the acceptance that child abuse occurs and that they as teachers have an involvement.

Confidence in School procedures
Question 6. I feel my school has a useful set of guidelines for staff to follow when dealing with suspected abuse.
Question 7. I am happy with the level of staff training for child protection which takes place in my school
Rationale - A measure of confidence in school procedures to deal with child abuse and protection was included in the light of many studies, of which that by Jones et al. (1987) is typical, which indicate teachers lack confidence.

Need for parental involvement
Question 8. I feel that regular contact with parents is very important for a child's progress and well being at school
Rationale - Researchers such as Elman (2000), indicate that parental involvement with all aspects of school is a good thing.

RESULTS
Hypothesis 1 The course participants would be more likely to use teaching materials which required less involvement/effort on their part than material which required more involvement/effort.

The hypothesis is supported for as can be seen in table 2 the most frequently used materials are the ones which in table 1 are seen to require the least effort. The number in the second column from the left of table 2 shows the ranking of each item on the effort scale. For the co-ordinators who attended the course the correlation between materials used and effort involved is .874 which is significant at the .033 level.
According to Sheehan & DuPrey (1996) this level shows a meaningful relationship between the two measures. For the co-ordinators who did not attend the course the correlation between the effort scale and the use made of materials is only slightly lower at .689. The teaching materials most used are those which can be given the children to use by themselves, those least used are the ones which include teacher involvement.

Table 2

Showing for use of materials in school comparisons between co-ordinators who did not attend the course (n=101) and those who did (n=273). The number to the right of the ‘materials used box’ shows the ranking of that item on the ‘effort scale’. 1 indicates smallest amount of effort and 6 the greatest amount.

<table>
<thead>
<tr>
<th>Materials used with most frequent at top. Number on left shows rank order on scale of effort</th>
<th>Mean rank order. Low score less effort</th>
<th>Mean scores</th>
<th>Standard deviations</th>
<th>Chi value df. = 1</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stranger Danger</td>
<td>2</td>
<td>.48</td>
<td>.50</td>
<td>.250</td>
<td>.617</td>
</tr>
<tr>
<td>Co-ordinators who did not attend the course</td>
<td>Co-ordinators who did attend</td>
<td>.47</td>
<td>.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kidscape</td>
<td>1</td>
<td>.19</td>
<td>.39</td>
<td>14.794</td>
<td>.000 *</td>
</tr>
<tr>
<td>Co-ordinators who did not attend the course</td>
<td>Co-ordinators who did attend</td>
<td>.29</td>
<td>.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Where to go for help</td>
<td>3</td>
<td>.12</td>
<td>3.2</td>
<td>43.299</td>
<td>.000 *</td>
</tr>
<tr>
<td>Co-ordinators who did not attend the course</td>
<td>Co-ordinators who did attend</td>
<td>.26</td>
<td>.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assertiveness training</td>
<td>4</td>
<td>.08</td>
<td>.28</td>
<td>14.339</td>
<td>.000 *</td>
</tr>
<tr>
<td>Co-ordinators who did not attend the course</td>
<td>Co-ordinators who did attend</td>
<td>.17</td>
<td>.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>How to deal with harassment</td>
<td>5</td>
<td>.05</td>
<td>.22</td>
<td>15.906</td>
<td>.000 *</td>
</tr>
<tr>
<td>Co-ordinators who did not attend the course</td>
<td>Co-ordinators who did attend</td>
<td>.12</td>
<td>.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teenscape</td>
<td>6</td>
<td>.02</td>
<td>.14</td>
<td>6.47</td>
<td>.011*</td>
</tr>
<tr>
<td>Co-ordinators who did not attend the course</td>
<td>Co-ordinators who did attend</td>
<td>.04</td>
<td>.20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The various materials used by the teachers raises doubt about how well they selected what they used. Looking at table 2, ‘stranger danger’, which was classed as easy to use is material which they can give to the children, without having to talk about it. Whether it should be the most used is doubtful as most of the threat to children comes from within the home, and not from strangers (Bowen 2000, Lewis & Creighton 1999 and Haapasalo & Aaltonen 1999).

Further in table 2 the comparison between the course participants and the control group of co-ordinators who did not attend the course provides evidence to support the impact of the child abuse course. Co-ordinators who attended the course reported a significantly greater use of five of the six types of teaching materials than did co-ordinators who did not attend.

Hypothesis 2. The course participants would be more likely to carry out activities which did not require the co-operation of others than activities which did require the co-operation of others.
This hypothesis, as can be seen in table 3, is supported for more activities which were under individual control had taken place in the schools attended by the course participants than co-ordinators who had not attended. These activities included revising school guidelines, borrowing child protection resources and attendance of staff at child protection courses other than INSETS. On face value it is encouraging for the course organisers to see that more activities were carried out by the sample who attended the training course than by those who did not. Using the ‘activities’ measure certainly makes the course appear to have been a success. This finding is in keeping with other studies which have shown the practical effects of training. Kneringer & Page (1999) found after training, the new behaviour which had been taught was continued for up to one year when the evaluation took place. Rosenblum (1999) reported in her survey that there was evidence early training in child abuse matters had a long lasting effect on clinical practitioners and had the effect of making them continue to take a greater interest in the topic than those who had not experienced early training. The result is quite heartening as an example of the positive effect of the course, for as McKenna (1994 pp 494-495) points out, the longer someone has been working the less likely they are to break old habits and adopt new methods. Some of the course participants had been teachers for many years.

Table 3
Showing comparisons between the co-ordinators who did not attend the course (n=101) and co-ordinators who did attend the course (n=273)

<table>
<thead>
<tr>
<th></th>
<th>Mean score</th>
<th>Standard deviations</th>
<th>t. score df. = 372</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Under own control</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Activities in school</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-ordinators who did not attend the course</td>
<td>0.94</td>
<td>1.06</td>
<td>3.02</td>
<td>.003 *</td>
</tr>
<tr>
<td>Co-ordinators who did attend the course</td>
<td>1.36</td>
<td>1.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Not under own control</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of referrals</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-ordinators who did not attend the course</td>
<td>1.65</td>
<td>.75</td>
<td>.356</td>
<td>.722</td>
</tr>
<tr>
<td>Co-ordinators who did attend the course</td>
<td>1.62</td>
<td>.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of INSETS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-ordinators who did not attend the course</td>
<td>2.74</td>
<td>2.25</td>
<td>1.77</td>
<td>.078</td>
</tr>
<tr>
<td>Co-ordinators who did attend the course</td>
<td>3.24</td>
<td>2.49</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A very different picture, however, emerges if activities which require co-operation are used as a measure. Schools with course participants were no more likely to make more referrals or to use INSETS for disseminating information about child protection than schools where members of the control group were based. If these two measures had been used as the criterion to evaluate the effectiveness of the course it would not have been considered successful.

**Why were INSETS not provided?**
Two hundred and thirty five out of three hundred and seventy respondents, of the questionnaires, failed to record that any child abuse/protection INSET had been carried out in their schools. Table 4 shows that the main reason given was they had other more urgent problems.

Hypothesis two was based on the assumption INSETS were not run because they were not under the control of the staff. Certainly from this questionnaire it is clear that lack of resources was not the cause. It would seem INSETS on procedures for dealing with child abuse were a low priority in the schools. Those who determine priority may well have been people other than the co-ordinators themselves. As a measure to judge the success of a course it would not seem to be an appropriate one as it depends on many factors outside the control of the course participants. It is evident from these findings in any evaluation of a course, whether the measures used are or are not under the direct control of the participants is a factor which can influence the validity of that measure.

Hypothesis 3. The course participants would display more positive attitudes towards their own role and school procedures for dealing with cases of child abuse than members of two control groups who had not attended the course.
Support for this hypothesis is mixed. Table 5 compares the four attitude and perceptual measures for the coordinators who attended the course and the coordinators who did not attend the course. The course had an effect on only three of the four measures. How this is interpreted, obviously, depends in the weight in terms of their significance each of the measures are. The relative importance of the measure where no difference occurred would need to be taken into account in any interpretation of the evaluation.

Table 4.

<table>
<thead>
<tr>
<th>Reasons in rank order with no 1 as the most often</th>
<th>Sample which attended course n=273 Mean score, standard deviation in brackets</th>
<th>Sample which did not attend course N=102 Mean score, standard deviation in brackets</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Other more urgent priorities</td>
<td>2.68 (2.46)</td>
<td>3.12 (2.34)</td>
</tr>
<tr>
<td>2. Financial constraints</td>
<td>1.06 (1.80)</td>
<td>1.75 (2.00)</td>
</tr>
<tr>
<td>3. No suitable course/resources</td>
<td>0.44 (1.14)</td>
<td>0.83 (1.53)</td>
</tr>
<tr>
<td>4. Staff cutbacks</td>
<td>0.34 (.95)</td>
<td>0.68 (1.29)</td>
</tr>
<tr>
<td>5. Unable to find someone to deliver it on the dates required</td>
<td>0.25 (.77)</td>
<td>0.52 (1.22)</td>
</tr>
</tbody>
</table>

Table 5.

<table>
<thead>
<tr>
<th>Measure and sub group</th>
<th>Mean scores</th>
<th>Standard deviation</th>
<th>t. score df = 372</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feelings about Knowledge of what to do:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-ordinators who did attend the course</td>
<td>3.26</td>
<td>.62</td>
<td></td>
<td>.002*</td>
</tr>
<tr>
<td>Co-ordinators who did not attend</td>
<td>3.03</td>
<td>.76</td>
<td>3.09</td>
<td>.002*</td>
</tr>
<tr>
<td>Acceptance abuse takes place:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-ordinators who did attend the course</td>
<td>13.12</td>
<td>1.87</td>
<td>2.20</td>
<td>.028*</td>
</tr>
<tr>
<td>Co-ordinators who did not attend</td>
<td>12.61</td>
<td>2.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confidence in school procedures:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-ordinators who did attend the course</td>
<td>5.76</td>
<td>1.28</td>
<td>3.52</td>
<td>.000*</td>
</tr>
<tr>
<td>Co-ordinators who did not attend</td>
<td>5.22</td>
<td>1.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Need for parent involvement:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-ordinators who did attend the course</td>
<td>3.82</td>
<td>.54</td>
<td>0.37</td>
<td>.712</td>
</tr>
<tr>
<td>Co-ordinators who did not attend</td>
<td>3.80</td>
<td>.58</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comparison with a ‘control’ group of teachers not directly involved with child abuse

Table 6 presents a more convincing picture of the effect of the course. When the course participants are compared to the qualified teachers who were not child abuse co-ordinators the former score more highly on all four measures. Using this comparison the course is shown to be remarkably successful. The results relating to the third hypothesis do raise a number of points about the way in which evaluations are compared. The attitude scales used produce very different results depending on which control group is used for comparison purposes.
Table 6
Showing differences between co-ordinators who did attend the course (n=273), and the qualified teachers sample (n=360), on the four attitude and perceptual measures.

<table>
<thead>
<tr>
<th>Measure and sub group</th>
<th>Mean scores</th>
<th>Standard deviation</th>
<th>t. score df = 631</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feelings about Knowledge of what to do:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-ordinators who attended the course</td>
<td>3.26</td>
<td>.62</td>
<td>8.06</td>
<td>.000*</td>
</tr>
<tr>
<td>Qualified teachers</td>
<td>2.80</td>
<td>.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acceptance abuse takes place:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-ordinators who attended the course</td>
<td>13.12</td>
<td>1.87</td>
<td>5.79</td>
<td>.000*</td>
</tr>
<tr>
<td>Qualified teachers</td>
<td>12.08</td>
<td>2.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confidence in school procedures:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-ordinators who attended the course</td>
<td>5.76</td>
<td>1.28</td>
<td>8.17</td>
<td>.000*</td>
</tr>
<tr>
<td>Qualified teachers</td>
<td>4.90</td>
<td>1.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Need for parent involvement:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-ordinators who attended the course</td>
<td>3.82</td>
<td>.54</td>
<td>7.84</td>
<td>.000*</td>
</tr>
<tr>
<td>Qualified teachers</td>
<td>3.36</td>
<td>.86</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

DISCUSSION
This study demonstrates activities carried out by course participants and used as measures to evaluate the effectiveness of a course may be a greater reflection of the preferences of the participants than they are of the course itself. This may have a marked significance on interpretations made from the results of any evaluation. It does seem, simply linking the performance of an activity with a judgement about the success of a course is not a very good idea. It is more likely to be a reflection of the effort required to produce the behaviour rather than on the merits of the course.

The second hypothesis shows how activities carried out in the school under the control of the individual co-ordinators, can demonstrate the impact of the course. On matters for which they had to liaise with others, for example making referrals or setting up INSETS, there was no evidence of an effect from the course. This highlights the need to ensure any measures used to evaluate a course are ones which are under the direct control of the participants on the course. Factors not under their control are less likely to be carried out and for that reason are unlikely to provide a helpful measure of the effectiveness of the course.

The third finding involves evaluations based on attitude change and here, unlike with the study by Tait & Purdie (2000), a marked effect was noticed. The effect did, however, vary depending on the control group used for comparison purposes, and this may explain the difference between the findings of this study and those of Tait & Purdie. When the course co-ordinators were compared to co-ordinators who had not attended the course they were seen to be more favourable on three of the four measures. When those who attended the course were compared with the qualified teachers who were not child abuse co-ordinators the impact of the course was even more convincing having an impact on all four attitude measures. Using only the co-ordinators who had not attended the course makes the course appear to have been less successful, than when the qualified teachers who were not involved in child abuse procedures were used as a control. The confounding variable here was possibly that the experience of doing the job, with the co-ordinators who were not ‘trained’, had imparted some of the attitudinal change obtained by those attending the course. This raises the importance of the need for the careful selection of a control sample.

The findings of this study show how measures designed to measure a variety of outcomes of a course can be influenced by a range of factors. There is a risk of accepting at face value the findings of any evaluation completed by participants on a course. There are a whole range of factors which determine how useful activity and attitude measures are as yardsticks of course effectiveness. The choice of groups used as ‘controls’ can also have an effect on whether a course is considered to be effective. Interpretations of the
effectiveness of evaluation reports by course participants need to be made with care as there are a number of factors having little relationship to how good the course actually was which may be influential.

Acknowledgement
This empirical study is integrated in the author’s own doctoral thesis.

Notes on contributor
Jenny Darby was Head of Science in an inner city Comprehensive. During a secondment she was responsible for a county wide training programme for head teachers. Recently she has been running teacher skills workshops for postgraduates at Loughborough University, where she obtained her doctorate.

REFERENCES


**Appendix**

Questionnaire edited into sections to relate to methods section of this article for clarity. Answer format included tick boxes which have been deleted for purposes of journal reproduction. Italic headings not included in version distributed to participants.

**Teaching materials measures**

Please tick any of the following from which materials are regularly included anywhere in the curriculum in your school

- Teenscape
- Stranger danger
- Kidscape
- Assertiveness training
- How to deal with harassment
- Where to go for help ie. Childline/Base51
- Other-please specify

**Activities carried out under the individual’s control**

Please tick any of the following carried out in your school during the past three years

- A revision of the school’s guidelines on dealing with suspected cases.
- Borrowing the child abuse/protection resources held by LEA professional centre.
- Attendance by any member of staff at a course concerning child abuse/protection out of school, but during school hours.

**Activities carried out NOT under the individual’s control**

How many referrals regarding suspected or actual abuse have been made to Social Services by your school during the last complete term? Please tick the appropriate number.

- None
- 1-3
- 4-6
- More than 6

Please tick any of the following carried out in your school during the past three years

- Whole day INSET on child abuse
- Half day INSET on child abuse
- Single session INSET on child abuse
- Staff meeting exclusively about child abuse

If no recent training for child protection/suspected abuse has taken place in your school, which do you believe of the following are the reasons? Please number in priority starting with number 1 as the most important.

- Financial constraints
- Other more urgent priorities
- Staff cutbacks
- No suitable courses/resources
- Other – please specify