

7 Phases 4 and 5: Methods

This section details the methods employed to achieve the following:

- Identify the design characteristics of risk assessment checklists which are most effective. This was achieved by comparing the ease of completion, subjective opinions and the effectiveness of two different designs of checklist-based risk assessment.
- Evaluate whether accompanying training in the use of the selected checklists is beneficial and whether benefits from training vary for the different designs of checklist. This was achieved by comparing the effectiveness of untrained and trained use of two different designs of checklist-based risk assessment.

The first method section (Section 7.1) presents the method employed to select companies and employees to participate in the project. Section 7.2 outlines the development of the two test checklists to be used in the project. Section 7.3 presents the method for the risk assessment trials and section 7.4 outlines the method for the longitudinal study.

7.1 Selection of participating companies

Information collected from 15 companies during the audit walk-through (phase 2 of the project) was used to select four companies for participation in the trials. All selected companies for inclusion in the trials needed to be similarly matched in terms of level of training, structure, work tasks and size. The following selection criteria were used to select four similar companies.

7.1.1 Level of training

Selected companies needed to provide a similar level of training in relation to awareness and risk assessment of MSDs to all potential participants (line managers, line leaders and line workers). This was ascertained from phase 2 (the walk-through audit) and results from the workplace questionnaire (Section 4.1.2).

7.1.2 Structure

Companies needed to consist of at least two separate sections from which two distinct groups of participants could be gathered to provide a control group (untrained in using checklist A or B) and a test group (to be provided with training in using checklist A or B). The study design used two discrete areas within a single company to enable better control for external factors, such as company budgets and attitudes towards health and safety, which may have an effect on the longitudinal part of the study in which comparisons are made between trained and untrained groups and the identification and implementation of solutions over a six-month period.

7.1.3 Work tasks

Companies needed to conduct work tasks that were similar in terms of physical and time demands. Tasks needed to be manually intensive, primarily involving the upper limbs, and repetitive in nature (but not manual handling). They also needed to be conducted in a production line setup.

7.1.4 Size

Selected companies needed to be similar in size in terms of overall company and number of employees working in each of the selected separate work areas.

7.2 Selection of participating employees

Each company was asked to provide eight to 10 people from each of the two distinct work areas in the company to participate in the study. It was requested that individuals should be selected who had either professionally or personally shown an interest in developing their risk assessment skills in relation to musculoskeletal problems. In addition, it was requested that all participants were able to understand, read and write English to a reasonable standard.

It was requested that each group should consist of one or two line managers, one to three supervisors (line leaders or team leaders) and six to eight production line workers (line operatives or line workers) from each area. The aim was to gain at least 16 participants from each company, providing a study cohort of approximately 64 participants. To ensure that each group from each company had similar training and understanding of MSDs, each participant was asked to complete a workplace questionnaire prior to the trials.

7.2.1 Workplace questionnaire

A workplace questionnaire was designed to probe participants' understanding of MSDs and their attitudes towards health and safety before completing the trials. The data were used to ensure that each group of participants comprised people with similar levels of understanding and training in MSDs in relation to their job position. In addition, the questionnaire provided the 'before' data for the longitudinal study. The workplace questionnaire was redistributed after a period of up to six months to ascertain whether, following the training, there had been any significant long term uptake of knowledge and/or changes in attitudes towards health and safety. A copy of the questionnaire is presented in Appendix D.

The participants completed the questionnaires a month before they attended the assessment trials. Due to the nature of some of the questions (eg regarding personal attitudes towards health and safety and feelings regarding the management of health and safety), the questionnaire was made anonymous to encourage true responses. In total 70 questionnaires were distributed.

7.3 Development of risk assessment checklists

Based on findings from the literature review conducted during phase 1, a standard risk assessment checklist for assessing ULD risk factors was selected.

The criteria from this checklist were then repackaged into two different formats to create two different test checklists (Checklists A and B), which differed in design features only, rather than in assessment criteria. Checklists A and B incorporated a range of design features previously identified and discussed in the literature review. Table 7.1 presents a summary of the design features included in each of the checklists and highlights the features where the two checklists differ. A detailed review and comparison of the two checklists for language, layout and so on is available from the authors on request.

The format used in Checklist A was based on a newly developed checklist by the HSE entitled 'Assessment of repetitive tasks of the upper limb' (ART). This is similar in design format to the now widely used MAC tool used for assessing manual handling tasks. The format incorporates a 'traffic light' rating scheme and provides numerical scores for each check item and a total overall risk score. The check items are presented in a tabular form with three or more response categories. A copy of Checklist A is presented in Appendix 4.

Checklist B is based on a more traditional list-type design in which each check item is presented in a list format with a 'yes/no' response as to whether the risk is present. Checklist B includes reference to a list of potential solution ideas and provides areas for the assessor to write more information about the probable cause and proposed solutions worthy of further investigation. A copy of Checklist B is presented in Appendix 4.

7.4 Risk assessment trials

7.4.1 Overview

The risk assessment trials were split into two sets of trials, referred to as trial 1 and trial 2. Trial 1 collected data on the use of checklists when used as a standalone tool (ie only using the accompanying written instructions in how to use and complete the checklists). Trial 1 was designed to collect data to:

- compare the effectiveness and ease of completing checklists A and B
- provide comparative qualitative data on participants' opinions of checklists A and B
- provide 'before' training data (completion of the checklists by untrained participants) to be compared to results gained from trial 2 (completion of the checklists by participants after receiving training).

Trial 2 collected data on the use of the checklists following a two-hour training session on their completion and use developed by an IOSH accredited trainer. Two of the companies were trained in the use of checklist A and two in checklist B. Data from the two trials were then compared.

Two groups from distinct areas in the same company participated in the trials (groups 1 and 2). All groups attended trial 1 but only group 1 from each company attended trial 2 (where training in the use of one of the checklists was given). Both groups from each company were encouraged to use only one of the checklists following the trials. Companies 1 and 2 were encouraged to use checklist A and companies 3 and 4 were encouraged to use checklist B. This was primarily for the purpose of the

		Checklist A	Checklist B
Format			
Flow chart			
List format		✓	✓
Multiple choice		✓	
Dichotomous			✓
Phrasing definitions			
Uses numbers to define joint angles		✓	
Uses words to describe joint angles			✓
Uses numbers to define frequency rates		✓	
Uses words to describe frequency rates			✓
Uses numbers to define duration			✓
Uses words to describe duration		✓	
Uses numbers to describe weight/force		✓	
Uses words to describe weights/force			✓
Visual aids			
Illustrations of postures			✓
Recording risk details			
Space for notes on reported problems			✓
Space and notes on risk /probable cause			✓
Ratings			
Means of rating individual items	Colour coding	✓	
	Symbol coding		
	Numerical	✓	
	Words, eg high, medium and low, good, satisfactory	✓	
Means of prioritising specific aspects of concern within a single task		✓	
Means of calculating an overall scores		✓	✓
Means of prioritising tasks for action		✓	✓
Controls and interventions			
Asks whether action is required			✓
Space for notes on potential actions			✓
Provides hints/suggestions for control actions	In checklist		✓
	In accompanying guidance		
	Provides reference to other sources of information	✓	✓

Table 7.1
Different design characteristics of Checklists A and B (features on which they differ are highlighted)

longitudinal study, to enable data to be collected that would allow comparison of the effects of completion of assessments and the generation of solutions within companies before training (group 2) and after training (group 1) after a period of up to six months (see Section 7.5).

7.4.2 Trial 1 – Untrained in the use of the checklists

Trial 1 consisted of two parts: the task assessments and a comparison questionnaire.

Task assessment

Four tasks were assessed by each company using either checklist A or B. Data were collected to investigate the comparative ease of completion and their effectiveness in terms of agreement with assessments completed by experts, which are referred to as the ‘model responses’.

Selection of tasks for assessment

Between four and six tasks from each participating company were initially recorded on video. One videoed task from each company was then selected for inclusion in the task assessment trials. The tasks were selected to present a range of different risk factors for assessment in the trials.

Model responses

Before the trials each of the tasks was assessed independently by three experts from ESRI using the checklists. These results provided the model responses for comparative measures to be made between participant assessments and the model correct responses. To ensure that the experts’ assessments provided the correct risk assessment result for each of the check items for assessment A and B; the experts were allowed to slow down and pause video footage. This allowed them to make detailed observation of the types and degrees of postures adopted while the task was being performed, and also to calculate repetition rates. Where the three expert results differed the experts met and jointly viewed the video footage and discussed each point of view regarding that particular check item. Where necessary freeze frames were taken and images enlarged to ascertain the correct assessment result. This process was conducted until an agreement on the correct response for that particular check item was achieved.

Trial procedure

The researcher gave a brief introduction about themselves, the university and the project. The slides and verbal protocol used are presented in Appendix G.

Prior to completing the assessments, the participants were provided with the appropriate assessment checklist and a two-sided A4 instruction sheet on how to complete the checklist. The participants were given 10 minutes to read the instructions and familiarise themselves with the assessment checklist. Each participant was then given up to 20 minutes to assess each task. A video of the task being assessed was played throughout the time given to complete the checklist. Each video also displayed a stopwatch to show the passing of time to assist in the calculation of task durations and cycle times.

All participants assessed three out of the four possible videoed tasks. Participants assessed two of the tasks using checklist A and one of the tasks using checklist B, or vice versa. The tasks presented to each participant for assessment were assigned such that each company overall completed assessments for all four tasks (Table 7.2) and that all participants gained experience in using both of the checklists. This enabled them to make comparisons on the performance and ease of use of the two different checklists at the end of trial in the comparison questionnaire.

Table 7.2
Which checklist
was used to assess
each task in each
company

	Company 1	Company 2	Company 3	Company 4
Task 1	A	A	B	B
Task 2	A	A	B	B
Task 3	B	B	A	A
Task 4	B	B	A	A

At the end of the trial participants from companies 1 and 2 were encouraged to consider using checklist A for the duration of the study, whereas companies 3 and 4 were encouraged to use checklist B.

Data collected

For each task assessed the following data were collected:

- for each check item – presence of risk factor (present/not present)
- for each risk factor – level of risk (red, amber, green): only for checklist A
- ease of completing each check item on a five-point scale, where 1 = very difficult and 5 = very easy
- total number of risk factors present
- overall risk rating of the whole task (low, medium or high)
- number of suggested improvements or changes provided to reduce the risks
- types of suggested improvements or changes made to reduce the risks.

Checklist A and B comparison questionnaire

Following the task assessments of trial 1, a questionnaire probing issues relating to the design and effectiveness of the two checklists was completed by all participants. The aim of the questionnaire was to provide quantitative and qualitative data on the ease of completing each checklist and to identify positive and negative design aspects of each checklist design used in the risk assessment trials. A copy of the questionnaire is presented in Appendix 5.

Figure 7.1 presents a flow diagram of Trial 1. The trial took approximately one and a half hours to complete. A copy of the full trial schedule and presentation is available from the authors on request.

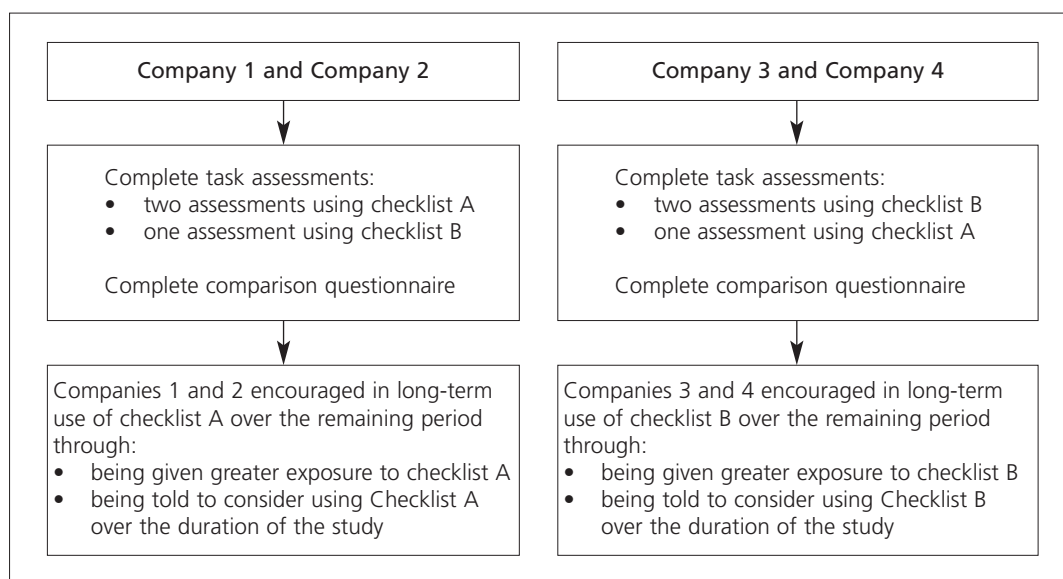


Figure 7.1
Trial 1 schematic

7.4.3 Trial 2 – Training in the use of the checklists

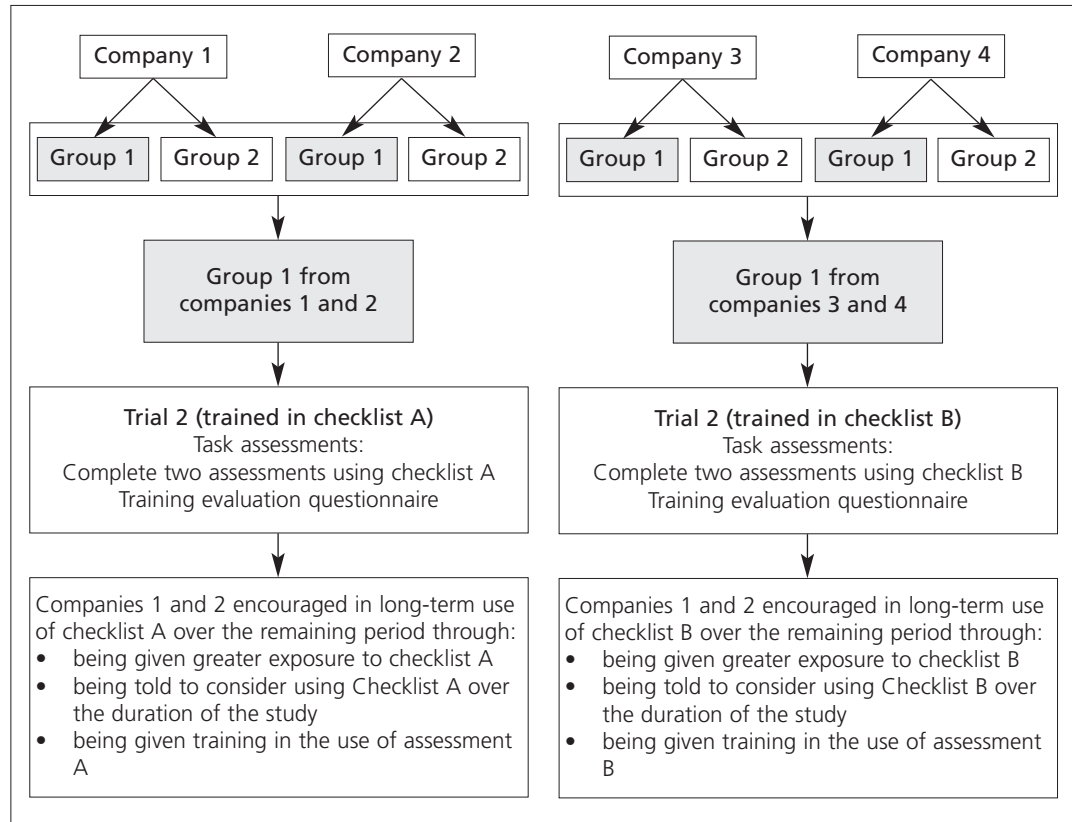
Trial 2 consisted of three parts: a training session, task assessments and a training evaluation questionnaire.

Training session

One of the selection criteria for participating companies was that each company should comprise at least two separate sections from which two distinct groups of participants could be gathered. This was to enable comparisons to be made for trial 2 and for the longitudinal study (phase 5 of the project) to investigate the difference between the groups that had received training and those that had not received training. Group 1 of each company provided the test group (participants that would be given training in the use of checklist A or B) and Group 2, the control group (participants untrained in using checklist A or B) (Figure 7.2).

Between five and six weeks after trial 1, group 1 from each company attended a two-hour training session in how to use assessment A or B by an IOSH accredited trainer and an ESRI lecturer. The

Figure 7.2
Trial 2 schematic



training session was first developed by the IOSH accredited trainer. The initial version of the training package and lesson plans are available from the authors on request. This training session was piloted with a group of eight employees from a company that was not participating in the study. Feedback from this session led to ESRI researchers further developing the course to:

- simplify the language used in the slides
- include more interactive segments
- expand the content regarding education about MSDs – what they are and what the risks are
- expand the content of how to complete an assessment using either checklist A or B
- include an example task where the group completes an assessment together with the trainer step by step.

A copy of the final course slides is available from the authors on request.

Task assessments

Following the training session, participants were asked to assess two tasks using the assessment checklist that they had just received training in. Companies 1 and 2 assessed the tasks using checklist A and companies 3 and 4 assessed the tasks using checklist B. Data were collected to investigate the ease of completion and effectiveness (in terms of agreement with the model responses) of each checklist, comparing their use before and after training. This was achieved by comparing results for tasks 1 and 2 from trial 1 to the results gained from trial 2.

Selection of tasks for assessment

All groups had assessed task 1 previously (without training), enabling a within-subject design for comparing the data. For task 2 the majority of participants had not previously assessed the task, so this was a between-subjects design in which group 1 (after training) results were compared to group 2 (without training) results. A more detailed explanation of the analysis of the data of trial 2 is presented in Section 8.3.

Model response

Each of the tasks had been assessed by three experts from ESRI using the checklists. This provided the model correct risk assessment results for assessment A and B to ascertain the degree to which the untrained and trained responses reflected the correct model response.

Procedure

Each participant assessed two tasks (tasks 1 and 2). The participants were given up to 20 minutes to assess each task. A video of the task being assessed was played throughout the time given to complete the checklist. Each video also displayed a stopwatch to show the passing of time to assist in the calculation of task durations and cycle times.

Data collected

For each assessment, the following data were collected:

- for each check item – presence of risk factor (present/not present)
- for each risk factor – level of risk (red, amber, green): only for checklist A
- ease of completing each check item on a five-point scale where 1 = very difficult and 5 = very easy
- total number of risk factors present
- overall risk rating of the whole task (low, medium or high)
- number of suggested improvements/changes provided to reduce the risks
- types of suggested improvements or changes made to reduce the risks.

Training evaluation questionnaire

A training evaluation questionnaire was developed to assess participants' understanding of ULDs and risk assessment after completing the training. The questionnaire was based on the IOSH trainer's standard feedback questionnaire with the addition of some of the questions from the workplace questionnaire to enable comparison with the data collected before the trials. All participants in trial 2 completed the questionnaire. A copy of the training evaluation questionnaire is presented in Appendix 6.

7.5 Longitudinal study

For the longitudinal part of the project participating companies were revisited up to six months after Trial 2 by an ESRI researcher to review progress.

The longitudinal study aimed to compare trained and untrained groups within each company and their identification and implementation of solutions following their exposure to the checklist tools.

The longitudinal study comprised two parts: completion of the workplace questionnaire (to compare against the initial results from the workplace questionnaire completed before trial 1) and an interview and walk-through with the health and safety officer of each group area of the site (group 1 and group 2 work areas).

Table 7.3 presents the activities that were conducted as part of the longitudinal study and Sections 7.5.1 and 7.5.2 outline the methods for recompletion of the workplace questionnaire and site visits.

Time	Activity
Prior to trials 1 and 2 (phase 2)	Interview and walk thorough with Health and Safety Officer.
Prior to trials 1 and 2	Trial participants complete workplace questionnaire
End of longitudinal period (up to six months after Trial 2)	Trial participants re-complete the workplace questionnaire. Interview and walk thorough with Health and Safety Officer.

Table 7.3

Time schedule for activities for the longitudinal part of the study (phase 5)

7.5.1 Recompletion of workplace questionnaire

Each participant from trial 1 completed the workplace questionnaire for a second time. The questionnaire results from before and after training for groups 1 and 2 were compared to identify whether training had resulted in increased confidence, what changes, if any, had occurred in attitudes towards health and safety, and whether the involvement of staff in identifying and reporting problems and proposing solutions had altered.

7.5.2 Interview and walk-through with health and safety manager

Interviews with health and safety officers were also conducted, including a walk-through of the group 1 and group 2 work areas. During the walk-through, the health and safety officer was asked to point out any changes to tasks that were being planned or had been made during the study period and to discuss:

- where the request for change had originated from (eg in response to reports made by workers, engineers, production staff, clients)
- why they had been made (eg to improve health and safety or production rates, in response to product changes, demands from external clients)
- what the changes were
- whether the changes had been effective in achieving the aforementioned goal(s)
- whether health and safety in relation to MSDs had been improved or worsened by the changes (expert appraisals were made of changed tasks in which expert ergonomists from ESRI gave a rating score for the practicalities and potential reduction in risks for each implemented intervention).

8 Phases 4 and 5: Data analysis

8.1 Workplace questionnaire

Quantitative and qualitative analyses of the data gained from the workplace questionnaire were conducted. Results relevant to the selection of participants are presented in Section 9.1.2. However, the majority of the results from the questionnaire are presented in Section 9.4.1 as part of the longitudinal phase of the project. In the longitudinal phase, the data collected from the questionnaire is compared to data collected from the re-completion of the same workplace questionnaire later. The changes in results gained from the groups that received training (group 1) and those that did not receive training (group 2) within each company over that period were compared.

8.2 Comparing checklists A and B (without training)

This section details the analysis conducted on six sets of data gathered from risk assessment trial 1 (Table 8.1). The sets of analyses were conducted to investigate the comparative effectiveness of checklists A and B. The results from the data analyses are presented in Section 9.2. In some instances (marked in Table 8.1) two levels of analysis have been conducted:

- level 1 provides an overview where data from all the assessed tasks have been combined
- level 2 shows the separate analysis of data gained from assessing each task.

The main body of this report presents the results and findings gained from level 1 analysis, which combines the data gathered from the assessments of all four tasks. Table 8.2 shows that across all four companies both checklists (A and B) were used to assess each of the four tasks. Combining all task data for each checklist reduces the effects which might be attributable to the type of task being assessed and/or differences which might arise from checklists being used by different participants.

Combining the result from all four tasks means that the pooled data for Checklist A will be directly comparable to data gathered using Checklist B, as the data for both checklists would then have been generated by the same participants from all four companies and across all four tasks.

The second level of analysis looks at each task separately. All the results from this level of analysis are presented in Appendix R and are only discussed in the main body of the report if the results are of particular interest or differ significantly from those gained by looking at the pooled task data (level 1 analysis).

8.2.1 Comparison questionnaire – participants opinions

A quantitative and qualitative analysis of the data gained from the comparison questionnaire was conducted. The quantitative data were appraised using largely descriptive statistics. The qualitative data were filtered and collated by hand with common themes recorded, grouped and counted. The outcome of this process was then collated and presented in comparative form. The prevalence of specific issues is therefore recognised and prioritised, while still retaining recognition of low incidence but high importance responses. The results are presented in Section 9.2.1.

8.2.2 Ease of completion of each check item

Level 1 – Analysis of all task data combined (Tasks 1 to 4)

Descriptive statistics of mean, mode and standard deviation of ratings for ‘ease of completion’ for each check item were calculated, enabling differences between Checklist A and B to be investigated. Furthermore, a between-subjects statistical analysis test was conducted on the pooled data (from all tasks) to compare the ‘ease of completion’ ratings for checklists A and B. The Mann-Whitney statistical test was applied to investigate whether any differences were statistically significant; the results are presented in Section 9.2.2. This was conducted first for all participants and then the analysis was repeated for participants split by job position.

Level 2 – Analysis of data for each task separately

A between-subjects statistical analysis test using Mann-Whitney was conducted to investigate whether differences between checklists A and B in ease of completing each check item were statistically significant for each separate task. These results are only discussed in the main body of the report if they are of particular interest or differ significantly from those gained by looking at the pooled task data. The full results are available from the authors on request.

Table 8.1
Types of analysis
conducted on six
sets of data

Data set	Level of analysis conducted	Type of analysis conducted
Comparison questionnaire	Not applicable	Quantitative and qualitative
Ease of completing Checklist A and B	1 (data from all tasks combined) 2 (data from each task separately)	Descriptive statistics Between-subjects design using Mann Whitney statistical test to compare companies 1 and 2 data with companies 3 and 4 for each checklist
Level of agreement between participants and model response	1 (data from all tasks combined) 2 (data from each task separately)	Percentage of participants that agreed with the Model response for each check item of each checklist.
Overall risk score (Total number of risk factors present)	1 (data from all tasks combined) 2 (data from each task separately)	A within-subjects statistical analysis test using Friedman A within-subjects statistical analysis test using Kruskal-Wallis A between-subject analysis using Mann Whitney.
Overall risk rating	1 (data from all tasks combined) 2 (data from each task separately)	Descriptive statistics Percentage of participants that agreed with the Model response rating
Suggestions for improvement	1 (data from all tasks combined)	Descriptive statistics

Table 8.2
Type of checklist
each company
used to assess
each task

	Checklist A		Checklist B	
Task 1	Company 1	Company 2	Company 3	Company 4
Task 2	Company 1	Company 2	Company 3	Company 4
Task 3	Company 3	Company 4	Company 1	Company 2
Task 4	Company 3	Company 4	Company 1	Company 2

8.2.3 Percentage agreement for each check item

Level 1 – Analysis of all data combined (tasks 1 to 4)

The percentage of participants who agreed with the model response for each check item was calculated for each checklist, each check item (1 to 13) and each task. The results of all four tasks were then combined to provide a single percentage value for each check item to enable a direct comparison between the overall performance of checklists A and B. The data were split and analysed in two different ways to investigate the level of agreement and the effect of:

- the type of checklist (A or B)
- the respondent's job position (team leader, line leaders, line managers, or line workers).

Level 2 – Analysis of data for each task separately

A comparison of each of the four different tasks was undertaken by looking at percentage agreement data gained for each task separately. These results are presented in Appendix R and are only discussed in the main body of the report if the results are of particular interest or differ significantly from those gained by looking at the pooled task data.

8.2.4 Total number of risk factors

The total number of risk factors reported as present for each completed checklist was calculated for each participant, for each task. The dependant variable was then calculated by subtracting the model response value (for the total number of risk factors present) from participant values (of the total

number of risk factors present). This was done for each task. All negative values were then changed to positive values to give a pure measure of the degree of discrepancy between participants' values and the model value for each task. These positive values provided the dependant variable. In the remainder of this report this measure is referred to as the absolute discrepancy value (ADV).

Level 1 – Analysis of all task data combined (tasks 1 to 4)

All discrepancy values from every participant for every task were first entered into a single database to investigate whether the type of task assessed had a significant effect on the ADVs. This was determined by applying a Kruskal Wallis statistical test. The findings showed that the type of task had no significant effect on the ADVs ($p = 0.153$) and therefore justified using the mean values from all tasks per participant to enable a within-subjects comparison of the data in the following set of analyses.

Within-subjects analysis

For each participant the discrepancy values for each task using checklist A were added and the mean value taken to give the mean ADV for checklist A. This was repeated for checklist B scores.

A within-subjects statistical test using Friedman was conducted to investigate whether differences between the mean ADVs for checklists A and B were statistically significant. To investigate the interactions of company, job position and checklist, the differences between mean ADV A and mean ADV B were calculated (this is referred to as the checklist difference). Tables 8.3 and 8.4 summarise the effects and interactions that were investigated and the statistical tests applied.

Effect investigated	Questions to be answered	Statistical test employed
Main effect of company on participant–expert scores before training	On average, before training, in what way do the participant–expert score differences change with company, regardless of the effect (if any) of other variables?	Kruskal-Wallis, using company as independent variable and mean discrepancy value as dependent variable
Main effect of job position on participant–expert scores before training	On average, before training, in what way do the participant–expert score differences change with job position, regardless of the effect (if any) of other variables?	Kruskal-Wallis, using job position as independent variable and mean discrepancy value as dependent variable
Main effect of checklist on participant–expert scores before training	On average, before training, in what way do the participant–expert score differences change with checklist, regardless of the effect (if any) of other variables?	Friedman, using mean discrepancy value of checklist A and mean discrepancy value of checklist B

Table 8.3

The effects investigated and statistical tests applied

Interaction investigated	Questions aimed at answering	Statistical test employed
Interaction of company and job position on participant–expert scores before training	On average, before training, to what extent is the effect of company on the participant–expert score differences modified by the effect of job position, regardless of the effect of other variables?	Kruskal-Wallis, using company and job position as independent variable, and mean discrepancy value as dependent variable
Interaction of job position and checklist on participant–expert scores before training	On average, before training, to what extent is the effect of job position on the participant–expert score differences modified by the effect of checklist, regardless of the effect of other variables?	Kruskal-Wallis, using job position and checklist as independent variable, and mean discrepancy value as dependent variable

Table 8.4

The interactions investigated and statistical tests applied

Level 2 – Analysis of data for each task separately

For each separate task a between-subjects statistical analysis test using Mann-Whitney was conducted to investigate and confirm whether differences in discrepancy values for checklists A and B were statistically significant. The results are presented in Appendix R and are only discussed in the main body of the report if the results are of particular interest or differ significantly from those gained by looking at the pooled task data.

8.2.5 Overall risk rating

Level 1 – Analysis of all task data combined (tasks 1 to 4)

The percentage of participants that agreed with the model response for overall risk rating was calculated for each checklist separately. The data were split and analysed in two different ways to investigate the level of agreement and the effect of:

- the type of checklist (A and B).
- the respondent's job position (line leader, line worker and so on).

Level 2 – Analysis of data for each task separately

The percentage of participants that agreed with the model response for overall risk rating was calculated for each checklist separately and for each task. These results are only discussed in the main body of the report if they are of particular interest or differ significantly from those gained by looking at the pooled task data. The full results are available from the authors on request.

8.2.6 Suggested improvements

Level 1 – Analysis of all task data combined (task 1 to 4)

Descriptive statistics of mean, mode and standard deviation of the number of suggestions were calculated, enabling differences between checklists A and B to be investigated.

8.3 Comparing trained and untrained users

This section outlines the analyses conducted on data gathered from trial 2 of the risk assessment trials (Table 8.5). The trial was a within-subjects design for Task 1, comparing results provided by group 1 before training to the results gained after training. Task 2 involved a between-subjects design comparing data from group 2 (untrained participants) to group 1 results (after receiving training). The results from the analyses are presented in Section 9.3. In this section, the data are analysed for each task separately. The following sets of analyses aim to investigate the comparative effectiveness of each checklist before and after training.

8.3.1 Ease of completion of each check item

Descriptive statistics of mean, mode and standard deviation for ratings of ease of completion for each check item were calculated, enabling differences between the completion of each checklist before and after training to be investigated. Furthermore, a within-subjects statistical analysis test was conducted. The Wilcoxon (for within-subjects) and the Mann-Whitney (for between-subjects) statistical tests were applied to investigate whether any differences in ease of completing each check item for checklists before and after training were statistically significant for tasks 1 and 2. The results are presented in Section 9.3.1.

8.3.2 Percentage agreement for each check item

Each checklist comprised 13 check items. The percentage of participants that agreed with the model response for each check item were calculated to enable a direct comparison of the performance of each checklist when completed by trained and untrained participants. The data from trial 2 for tasks 1 and 2 (trained) were compared to data gained from trial 1 for tasks 1 and 2 (untrained).

8.3.3 Total number of risk factors

The total number of risk factors reported by each participant was calculated for each task and the discrepancy value then calculated. A within-subjects analysis was conducted on data gained for task 1 (Table 8.6) and a between-subjects analysis was conducted on data gathered for task 2 (Table 8.7). These statistical analyses compared the ADVs from trial 1 (no training) and trial 2 (with training) for each type of checklist for each of the two tasks. Table 8.8 presents a summary of the type of statistical tests employed.

Data set	Data analysed	Type of analysis conducted
Ease of completing Checklist A and B	Task 1 Group 1 data	Descriptive statistics Within-subjects design comparing group 1 before training with group 1 after training
	Task 2 Groups 1 and 2 data	Descriptive statistics Between-subjects design comparing group 2 before training with group 1 after training
Level of agreement between participants and model response	Task 1 Group 1 data	Percentage of participants from group 1 that agreed with the model response for each check item comparing before and after training results
	Task 2 Groups 1 and 2 data	Percentage of participants from group 1 that agreed with the model response for each check item comparing before (group 2) and after training (group 1) results
Overall risk	Task 1 Group 1 data	Descriptive statistics Within-subjects design comparing group 1 before training with group 1 after training
	Task 2 Group 1 and 2 data	Between-subjects design comparing group 2 before training with group 1 after training
Overall risk rating	Task 1 Group 1 data (all companies)	Descriptive statistics Percentage of participants that agreed with the model, comparing group 1 before and after training results
	Task 2 Group 1 and 2 data	Descriptive statistics Percentage of participants that agreed with the model, comparing before (group 2) and after training (group 1) results
Suggestions for improvement	Data from each task separately	Descriptive statistics

Table 8.5

Analyses conducted on data gathered from trial 2

8.3.4 Overall risk rating

The percentage of participants that agreed with the model response for overall risk rating was calculated for checklists A and B before and after training. The data was split and analysed in two different ways to investigate the level of agreement and the effect of:

- the type of checklist (A or B)
- the respondent's job position (line leader, line worker and so on).

8.3.5 Suggested improvements

Descriptive statistics of mean, mode and standard deviation for the number of suggestions were calculated for before and after training, enabling differences between each checklist before and after training to be compared. Respondents' descriptions of the types of change suggested are also reported for tasks 1 and 2, split by job position and type of checklist used.

8.3.6 Training evaluation questionnaire

A simple analysis was undertaken of the training questionnaire in order to establish the participants' views on the effectiveness of the training that they had received. This was not intended to be a comprehensive evaluation, since the participants' experiential backgrounds varied, which would significantly colour their stated preferences. Additionally, while qualitative feedback was welcomed, the expertise of the individuals would not necessarily qualify them to recommend good or bad elements.

Accordingly, while it was felt that there was value in assessing the training experience, this aspect of the work was of a lower priority. Its main aim was to inform further studies on the content and style of presentations so that they may appeal to participants rather than necessarily increase their effectiveness.

Table 8.6 (left)
 Type of checklist used by participant groups for assessing task 1, and type of analysis carried

Table 8.7 (right)
 Type of checklist used by participant groups for assessing task 2, and type of analysis carried

Company	1		2		3		4	
Checklist	A		A		B		B	
Trial	1. Before training	2. After training	1. Before training	2. After training	1. Before training	2. After training	1. Before training	2. After training
Participant group	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1
Task 1	Compare data (Within-subjects Wilcoxon)		Compare data (Within-subjects Wilcoxon)		Compare data (Within-subjects Wilcoxon)		Compare data (Within-subjects Wilcoxon)	

Company	1		2		3		4	
Checklist	A		A		B		B	
Trial	1. Before training	2. After training	1. Before training	2. After training	1. Before training	2. After training	1. Before training	2. After training
Participant group	Group 2	Group 1	Group 2	Group 1	Group 2	Group 1	Group 2	Group 1
Task 2	Compare data (Between-subjects Mann Whitney)		Compare data (Between-subjects Mann Whitney)		Compare data (Between-subjects Mann Whitney)		Compare data (Between-subjects Mann Whitney)	

Effect or interaction investigated	Questions aimed at answering	Statistical test employed
1. Main effect of company on participant–expert scores for group 1 after training	On average, for group 1, after training, in what way do the participant–expert score differences change with company, regardless of the effect (if any) of other variables?	Kruskal-Wallis using Co as IV and AT as DV
2. Main effect of checklist on participant–expert scores for group 1 after training	On average, for group 1, after training, in what way do the participant–expert score differences change with checklist, regardless of the effect (if any) of other variables?	Kruskal-Wallis using CL as IV and AT as DV
3. Interaction of company and checklist on participant–expert scores for group 1 after training	On average, for group 1, after training, to what extent is the effect of company on the participant–expert score differences modified by the effect of checklist difference, regardless of the effect of other variables?	Kruskal-Wallis using Co_CL as IV and AT as DV
4. Main effect of training on participant–expert scores	On average, after training, in what way do the participant–expert score differences change with Training, regardless of the effect (if any) of other variables?	Freidman's analysis of variance or Wilcoxon signed ranks test for task 1 checklist A
4. Main effect of training on participant–expert scores	On average, after training, in what way do the participant–expert score differences change with Training, regardless of the effect (if any) of other variables?	Freidman's analysis of variance or Wilcoxon signed ranks test for task 1 checklist B
5. Main effect of training on participant–expert scores	On average, after training, in what way do the participant–expert score differences change with training, regardless of the effect (if any) of other variables?	Mann Whitney (between-subjects) for task 2 checklist A
5. Main effect of training on participant–expert scores	On average, after training, in what way do the participant–expert score differences change with training, regardless of the effect (if any) of other variables?	Mann Whitney (between-subjects) for task 2 checklist B

Table 8.8
Within-subjects data analysis of task 2 (all companies data) and between-subjects data analysis of task 2 (companies 1 and 4 data)

9 Phases 4 and 5: Results

This section presents the results under the following headings.

- participants – companies and employees
- trial 1
- trial 2
- longitudinal study

9.1 Participants

9.1.1 Companies

Four companies previously visited in phase 2 of the study were selected to participate in the trials. The companies were selected because they were similar in terms of:

- level of training in MSDs
- comprising two distinct sections or areas
- type of work tasks undertaken
- size – number of employees working in each participating work area.

Tables 9.1 and 9.2 present a summary description of each of the selected participating companies.

9.1.2 Employees

Between nine and 20 employees from each company participated in the project, providing a study cohort of 63 participants (Table 9.3).

All participants from each company had received similar company training and education and had similar awareness of MSDs. This was ascertained from the results from the workplace questionnaire (discussed below) and interviews with health and safety officer/managers (summary presented in Table 9.2).

Workplace questionnaire

In total 48 participants completed the workplace questionnaire (Table 9.4).

Results from the questionnaire showed that over 85 per cent of respondents from companies 1, 2 and 4 had heard of RSI or MSDs (Table 9.5). Company 3 had a significantly lower percentage of respondents reporting that they had heard of either. However it should be noted that only six of the 15 participants from this company completed the workplace questionnaire, so the results are not reflective of the entire participant group.

Of the participants that reported that they had heard of MSDs or RSI, between 44 and 60 per cent of these stated that they had heard of MSDs through work, although only between 0 and 25 per cent said that this was from attending a work training course (Table 9.6). This is in contrast to reports made by the health and safety officer or manager of each company regarding training and education given to staff on MSDs or RSI. Health and safety officers and managers from all companies reported that training was provided to staff either during risk assessment training sessions or general work task training (provided to line supervisors, leaders and managers) and through induction courses (provided to line workers) (Table 9.7). The largest percentage of participants from all companies reported that they had heard about MSDs or RSI from television and magazines (Table 9.6).

Question 11 of the workplace questionnaire investigated people's understanding and knowledge of MSDs, and asked respondents to list up to six risk factors which may lead to musculoskeletal problems or RSI. Table 9.8 shows that similar numbers of risk factors were reported from respondents from all companies except company 3, in which a higher percentage (50 per cent) reported not knowing any risks or causes (Table 9.8 and Table 9.1).

Table 9.1
Summary
descriptions of
each participating
company

	Company 1	Company 2	Company 3	Company 4
	The company imports flowers from around the world, makes them into bouquets and packs them into boxes to distribute across the UK. The site employs 255 workers and comprises three separate areas or factories. Employees are paid by salary (not piece rate).	The company is a laboratory that tests soil, water and asbestos samples. The company employs approximately 110 people. Employees are paid by salary (not piece rate). There are five areas of work: cold store, sample preparation, samples reception and logging in, fume cabinets, and analysis.	The company prepares and packages salad. The company is part of a larger group but each company within this group acts independently. Therefore there are no common health and safety procedures prescribed for all the member sites or companies. The company employs 180 workers. Employees are paid by salary (not piece rate).	The company manufactures a range of desserts although the predominant product is cake. The company is part of a large group, which has a further 15 food production sites across the UK. The health and safety manager for this site is only responsible for this site. The site employs 800 staff across three shifts (266 per shift).
Groups	1 (factory 1) approx 30–40 workers 2 (factory 2) approx 30–40 workers	1 (site 1) approx 30 workers 2 (site 2) approx 30 workers	1 (low risk) 25–30 workers 2 (packing) approx 20 workers	1 (zone 1, line 3) approx 20 workers 2 (zone 2, line 1) approx 20 workers
Description of distinct groups	- Physically separate buildings but on the same site. - Workers from the two factories have separate rest areas. - Workers do not swap between different factories. Both factories complete very similar work activities but are split based on the different companies that they supply to.	- Physically separate buildings at different locations in the UK. - Workers do not swap between different sites. - Both laboratories complete identical work activities. Both laboratories are under the supervision of the same Health and Safety Officer and have the same health and safety policies and training programmes.	- Physically separate areas within the same large building. - Cannot walk from one area to the next without pass/clearance and going through appropriate clothing and cleaning regimes. - Workers from the two areas have separate rest areas. - Workers very rarely swap to work within the different work areas. Both areas complete different work activities but they are similar in nature, physical demands, postures and cycle times.	- Physically separate areas within the same large building. - Cannot walk from one area to the next without pass/clearance and going through appropriate clothing and cleaning regimes. - Workers very rarely swap to work within the different work areas. Both areas complete different work activities but they are similar in nature, physical demands, postures and cycle times.

Table 9.2
Summary
descriptions of the
tasks carried out
by each
participating
company

	Company 1	Company 2	Company 3	Company 4
Type of tasks conducted	Both groups conduct similar tasks in similar setups. Production line setup. Line feeders feed flower stems to line workers. Line workers conduct one of the following tasks along the line: • bunch formers • hand tying • sleeving the bunches • labelling • bottom and length check, trim stems • packing – make box All tasks are highly repetitive and of short cycle time. Machine paced.	Both groups conduct similar tasks in similar setups. Employees work at shared sit-stand workstations (ie each day they could be working at a different work stations conducting input tasks or preparation tasks). Inputting tasks include: • lifting and carrying sample trays to workstations • lifting out individual sample containers from the trays • inputting data into a computer • returning samples into trays • lifting trays back to reception to be placed into storage. Preparation tasks include: • lifting out individual sample containers from the trays • unscrewing sample containers • sifting samples • weighing samples • inputting data into a computer • returning samples into the trays. All tasks are highly repetitive and of short cycle time.	Both groups conduct similar tasks in similar setups. <i>Low risk</i> There are three lines in preparation. The preparation tasks conducted on all the lines are similar and include: • picking up individual produce from the conveyor • taking out the core of lettuces using a hand-held knife • removing dead leaves. Each item is also visually inspected for foreign bodies. The produce is then placed back onto a different conveyor. Repetition rate are high and machine paced. <i>Packing</i> There are three lines in packing. The packing tasks conducted on all the lines are similar and include: • picking up individual packaged produce from the rotating conveyor • checking that labels are correct • placing packaged produce into boxes. Both groups conduct similar tasks in similar setups.	<i>Zone 1</i> Each line produces a different type of cake. For each line the production process is very similar and includes the following tasks: • greasing cake tins • lifting filled cake tins and feeding into the oven • applying filling to one side of the sponge (palette knife or hand) • placing top sponge onto bottom sponge • coating cake (icing etc) • applying sprinkles/decoration • piping • inspection • removing from conveyor <i>Zone 2</i> Each line produces a type of prepackaged dessert. This includes placing product components into plastic trays or packages. Different workers along the line each have a different task, adding a different product component. Repetition rates are high and tasks are machine paced.

	Company				Total participants
	Company 1	Company 2	Company 3	Company 4	
No. of participants	20	19	15	9	63

Table 9.3
Number of employees in each company who took part in the trials

	Company				Total completed questionnaires
	Company 1	Company 2	Company 3	Company 4	
No. of completed questionnaires	15	18	6	9	48

Table 9.4
Number of completed workplace questionnaires

	Percentage of respondents			
	Company 1	Company 2	Company 3	Company 4
Yes, I have heard of MSDs	7	22	33	0
Yes, I have heard of RSI	73	22	33	44
Yes, I have heard of both MSDs and RSI	20	44	0	56
No, I have not heard of either	0	11	33	0

Table 9.5
Summary of training levels and awareness of MSDs among participants in each company

	Percentage of respondents			
	Company 1	Company 2	Company 3	Company 4
Television	60	63	7	33
Radio	7	31	25	11
Books	7	44	0	0
Magazines	53	56	25	33
Websites	0	25	0	0
Work	60	50	50	44
Training course	20	25	0	11
Doctor	27	19	25	44
Physiotherapist	13	19	0	22
Other	7	19	0	22

Table 9.6
How respondents had heard of MSDs or RSI

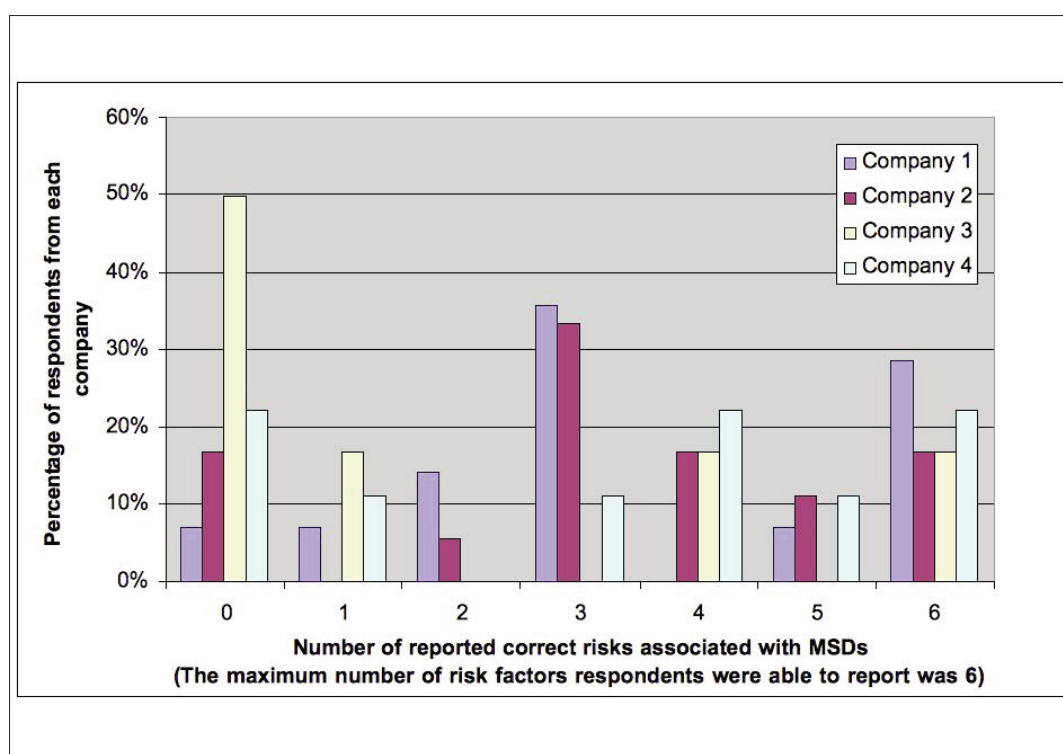
Table 9.7
Results of audit walk-through: training provided by each company as reported by health and safety managers

Company 1	Company 2	Company 3	Company 4
Potential participants			
- Team leader or supervisor - Line leader - Operative or line worker	- Laboratory manager - Team leader - Laboratory operative	- Line supervisor - Cell leader - Line worker	- Line leader - Team leader - Bakery operative or line worker
Level of training previously provided in MSDs			
Team leader Receive training in conducting general risk assessment. This is a package designed by the Health and Safety Manager. Undocumented training in the hazards and what to look for regarding MSDs and manual handling, e.g. poor postures.	Laboratory manager Receive training from the Health and Safety Manager about musculoskeletal disorders symptoms, risk factors and general good practice.	Line supervisor Attend an induction in which specific risks associated with their work tasks are discussed and explained including MSDs.	Line supervisor Receive training from the Health and Safety Manager about MSD symptoms, risk factors and general good practice.
Line leader Receive a 2.5 hour induction training covering all aspects of the job. This includes a manual handling video. They also receive a handout about ULDs to inform them of the symptoms and increase awareness of the issues. This forms part of the induction package.	Team leader Receive training from the Health and Safety Manager about musculoskeletal disorders symptoms, risk factors and general good practice.	Cell leader Attend an induction in which specific risks associated with their work tasks are discussed and explained.	Team leader Receive training from the Health and Safety Manager about MSD symptoms, risk factors and general good practice.
Line worker Receive a handout about ULDs to inform them of the symptoms and increase awareness of the issues. This forms part of the induction package.	Laboratory operative Attend an induction in which specific risks associated with their work tasks (including ULDs) are discussed and explained.	Line worker Attend an induction in which specific risks associated with their work tasks are discussed and explained.	Line worker Attend an induction in which specific risks associated with their work tasks (including ULDs) are discussed and explained. Training on the importance of micro pauses and hand exercises given to all line staff.

	Company 1	Company 2	Company 3	Company 4
Mean	3.5	3.3	1.8	3.2
Standard deviation	1.9	1.9	2.6	2.4
Minimum	0	0	0	0
Maximum	6	6	6	6

Table 9.8

Descriptive statistics for the number of risks or causes out of a potential total of six provided by each company

**Figure 9.1**

Percentage of respondents from each company and the number of correct risks reported

9.2 Trial 1 – Comparing checklists A and B (without training)

This section presents the results for investigating the comparative effectiveness of Checklist A and Checklist B. The results are presented under six headings:

- comparison questionnaire – participants' opinions
- ease of completion of each check item
- percentage agreement of each check item between participants and the model response
- total number of risk factors
- overall risk rating
- suggestions for improvement.

9.2.1 Comparison questionnaire – participants' opinions

Respondents

In total 57 participants completed the comparison questionnaire. Sixteen were team leaders, line leaders or line managers and 41 were line workers or line operatives. Table 9.9 presents the number of respondents from each company.

Preferred checklist

The majority of respondents reported that they would prefer to use checklist A (72 per cent) to assess their workplace rather than checklist B. However, when split by job position a significant difference in preferences emerged, with 63 per cent of team leaders, line leaders and line managers preferring checklist B, compared to 15 per cent of all line workers. Conversely, the majority of line workers (85 per cent) stated that they would prefer to use checklist A.

Table 9.9
Number of respondents who completed the comparison questionnaire

Company	Work position		Number of respondents
	Team leader, line leader, line manager	Line worker, operative	
1	6	10	16
2	4	14	18
3	3	12	15
4	3	5	8
Total	16	41	57

Similar results were found regarding responses to which of the two checklists respondents thought would be the best to help their company reduce the risks of musculoskeletal problems in their workplace. Sixty-one per cent of all respondents chose checklist A and 39 per cent checklist B. However, when split by job position it was found that the majority (69 per cent) of the team leaders, line leaders and line managers reported that checklist B would be the most helpful whereas most (73 per cent) of the line workers considered checklist A to be better.

The reasons for the differences in preference were explored further in questions 12 and 13 of the questionnaire, which asked respondents to list the two things they liked and disliked most about each checklist. Tables 9.10–9.13 present summaries of the responses, split by respondents' job position.

Positive comments – Checklist A

In total 51 (45 per cent) comments out of 114 received regarding Checklist A were positive. When split by job position, it was found that from a total of 32 comments from team leaders, line leaders and line managers, 13 (41 per cent) were positive, while from a total of 82 comments from line workers, 38 (46 per cent) were positive. Table 9.10 shows all the comments received. The following positive aspects were reported by the greatest number of respondents from each job position group:

- clear definitions, easy to understand, or explained things better than B
- use of colour coding
- easy to follow and/or use
- number/scoring.

Positive comments – Checklist B

In total 41 (36 per cent) comments out of 114 received regarding Checklist B were positive. When split by job position, it was found that from a total of 32 comments from team leaders, line leaders and line managers, 12 (38 per cent) were positive, while from a total of 82 comments from line workers, 29 (35 per cent) were positive. Table 9.11 shows all the comments received. The following positive aspects were reported by the greatest number of respondents from each job position group:

- illustrations
- presentation of possible solutions that you could tick if appropriate
- space to write and describe tasks and any particular problems in detail
- the number of choices (yes/no) made it easy to decide between them
- easy to follow and/or use.

Negative comments – Checklist A

In total 24 (21 per cent) comments out of 114 received regarding Checklist A were negative. When split by job position, it was found that from a total of 32 comments from team leaders, line leaders and line managers, 10 (31 per cent) were negative, while from a total of 82 comments from line workers, 14 (17 per cent) were negative. Table 9.12 shows all the comments received. The following negative aspects were reported by the greatest number of respondents from each job position group:

- no space to describe problems or make comments; another person would not be able to tell much about the task just from reading a completed assessment
- too many options
- no illustrations.

	Team leaders etc	Line workers	Total respondents
Clear definitions or Easy to understand or Explained things better than B	2	13	15
Use of colour coding	2	9	11
Easy/simple to follow or use	4	7	11
Number or scoring	3	5	8
Format seemed logical	1	1	2
More options to choose from, making it easier to complete	1	2	3
Automatically assigned a level of risk (high, medium, low)	0	1	1
Total number of positive comments	13	38	51

Table 9.10
Positive comments
on checklist A

	Team leaders etc	Line workers	Total respondents
Illustrations	3	7	10
Presentation of possible solutions that you could tick if appropriate. Made it easier to identify solutions	4	4	8
Space to write and describe tasks in detail and any particular problems	2	4	6
Number of choices (yes/no) easy to decide between	0	6	6
Easy/simple to follow or use	1	3	6
More in-depth explanations than A	0	4	1
Number or scoring – easy to score	1	0	1
Clear and concise	1	0	1
Layout of questions	0	1	1
Total number of positive comments	12	29	41

Table 9.11
Positive comments
on checklist B

Negative comments – Checklist B

In total 37 (32 per cent) comments out of 114 received regarding Checklist B were negative. When split by job position, it was found that from a total of 32 comments from team leaders, line leaders and line managers, seven (22 per cent) were negative, while from a total of 82 comments from line workers, 30 (37 per cent) were negative. Table 9.13 shows all the comments received. The following negative aspects were reported by the greatest number of respondents from each job position group:

- the overall risk level is not calculated, but is rather just left to be rated by the individual assessor; this is open to interpretation
- required more time
- not easy to follow
- not short or simple to complete; too wordy
- the yes/no questions were easy to complete but did not inform much about the task

Table 9.12
Negative
comments on
checklist A

	Team leaders etc	Line workers	Total respondents
No space to describe problem or make comments. Another person would not be able to tell much about the task just from reading a completed assessment	3	2	5
Too many options	2	1	3
No illustrations	1	2	3
Too much reading for each level of risk – ie too wordy	1	1	2
No list of improvements	0	2	2
The colour coding meant nothing in the final scheme of things	0	2	2
Sometimes no middle ground option	1	0	1
Use of percentages for time/duration	1	0	1
Different grades are not detailed enough for example, most of the time, part of the time, more than half of the time.	0	1	1
Each point had different scores for green, amber and red	0	1	1
People with colour vision problems may find it difficult to read and follow	0	1	1
Time consuming	1	0	1
No space next to each factor to describe possible solutions – only one space at end	0	1	1
Total number of negative comments	10	14	24

- having to describe causes of problems
- did not really understand the questions.

Ease of comprehension and use

The respondents were first asked to state which of the two checklists they found easier to use. Overall, a majority preferred checklist A (63 to 46 per cent), but there was a noticeably different response from each worker group. The team leaders were split 50-50 in their preference for each checklist, while more line workers preferred checklist A (76 to 24 per cent).

Respondents were then asked to rate the checklists on the following criteria, giving their answers on a five-point scale where 1 = least confident, effective or easy and 5 = most confident, effective or easy:

- confidence in using the checklists
- confidence that they had assessed the risks correctly
- effectiveness in identifying the tasks as high, medium or low risk
- effectiveness in identifying the causes of the problems
- effectiveness in identifying the changes needed to reduce the risks
- ease of following the instructions on how to use the assessment tools.

Table 9.14 shows the percentages of all respondents, and of those in the manager and worker groups separately, who answered these questions with a rating of 4 or 5.

	Team leaders etc	Line workers	Total respondents
The overall risk level is not calculated rather just left to be rated by the individual assessor. Open to interpretation and self opinion	2	7	9
Required more time	1	4	5
Not easy to follow	0	4	4
Not short or simple to complete. Too wordy	0	4	4
The yes/no questions were easy to complete but did not inform much about the task.	0	3	3
Having to describe causes of problems	0	3	3
Did not really understand the questions.	0	3	3
Not enough information for each risk factor	1	0	1
Too much writing required	1	0	1
Improvement list not always relevant to the task being assessed.	1	0	1
Unclear exactly what the risk was	1	0	1
Lack of any individual risk level for each risk factor	0	1	1
Not enough background info to make suggestions	0	1	1
Total number of negative comments made	7	30	37

Table 9.13
Negative
comments on
checklist B

Respondents' views on the need for additional information and means of providing that additional information

Respondents were asked to circle the response that best described their level of agreement with each of the statements A–G below:

- A I would have liked more background information about the risk factors.
- B I would have liked more information about how to complete the risk assessments.
- C I would have liked more information about possible control actions or changes to make to reduce the risks.
- D I think face-to-face training in the use of the assessments would be more useful than following written instructions.
- E I think following written instructions would be more useful than attending a face-to-face training session.
- F I thought that the written instructions were sufficient to conduct the assessments.
- G I would have liked to have gone through some example assessments with a trainer.

The full results are shown in Table 9.15 but the key points are summarised below.

- 64 per cent of all participants would have liked more background information about the risk factors.
- There were mixed views on whether more information was needed about how to complete the risk assessments, with 46 per cent agreeing that they would have liked more information and 30 per cent reporting that they did not require any more information than that which was provided.
- Over half of the respondents (62 per cent) agreed that they would have liked more information about possible control actions or changes to reduce the risks.
- 62 per cent of the respondents agreed that they would have liked face-to-face training rather than just written instruction and only 11 per cent disagreed with this. Over half (57 per cent) agreed that they would have liked to have gone through some example assessment with a trainer.

Table 9.14
Percentage of respondents giving a rating of 4 or 5 to the criteria listed

	% respondents giving a rating of 4 or 5					
	All respondents		Managers		Workers	
Checklist	A	B	A	B	A	B
Confidence in using the checklist	74	46	63	44	78	47
Confidence in assessing tasks correctly	49	33	50	63	48	22
Effectiveness in identifying risk as high, medium or low	69	35	50	62	77	25
Effectiveness in identifying causes of problems	60	38	58	50	63	33
Effectiveness in identifying changes to reduce risk	42	49	19	57	51	47
Perceived ease of use	68	32	50	50	76	24
Ease of following instructions on using assessment tools	63	46	32	25	76	53

Table 9.15
Percentage of respondents agreeing or disagreeing with statements A–G

Statements	% respondents				
	Strongly agree	Agree	Neither	Disagree	Strongly disagree
A. I would have liked more background information about the risk factors	7	57	30	5	0
B. I would have liked more information about how to complete the risk assessments	7	39	25	26	4
C. I would have liked more information about possible control actions/changes to make to reduce the risks	13	49	27	11	0
D. I think face-to-face training in the use of the assessment would be more useful than following written instructions	18	44	28	11	0
E. I think following written instructions would be more useful than attending a face-to-face training session	0	18	33	42	7
F. I thought that the written instructions were sufficient to conduct the assessments	2	61	28	7	2
G. I would have liked to have gone through some example assessments with a trainer	7	47	30	16	0

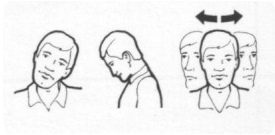
- However, 63 per cent of the respondents did agree that the written instructions that they had received with each checklist were sufficient to conduct the assessments.

Preference for risk descriptors

Question 12 was designed to investigate the respondents' preferences for the numerical or verbal descriptors for repetition, force, time and postures, and also the use of illustrations. Each respondent was asked to pick one of two different descriptions (term 1 or 2) which described the same risk. Table 9.16 presents the percentage of respondents and their preference to either term 1 or 2 for descriptions A to E.

Significant differences were only observed in respondents' responses to descriptions D and E. Notably, these were the descriptors that included an image. In other respects this section did not reveal anything significant regarding preferences of descriptive terms such as comparing word descriptions to numerical terms, or descriptions of time as a percentage compared to a word description.

Table 9.16
Respondents' preferences for terms 1 or 2 for descriptions A–E

Description	Term 1	Term 2
A	Does the task involve similar motion patterns being repeated frequently	Similar motion patterns are repeated more than 11–20 times per minute
Respondents' preference (% respondents)	44.0	55.6
Used in checklist	B	A
B	Moderate force (1–4 kg) or strong force (more than 4kg) is exerted	Moderate or strong force is exerted
Respondents' preference (% respondents)	54.4	45.6
Used in checklist	A	B
C	Does the task involve holding the neck bent or twisted more than 15% of the time?	Does the task involve holding the neck bent or twisted a part of the time?
Respondents' preference (% respondents)	49.1	50.9
Used in checklist	B	A
D	Does the task involve holding the neck bent or twisted? 	Does the task involve holding the neck bent or twisted (more than 15 degrees relative to the upright and forward facing position)?
Respondents' preference (% respondents)	75.0	25.0
Used in checklist	B	A
E	The back is bent forward, sideways or twisted (more than 20 degrees from upright forward facing position)	The back is bent forward, sideways or twisted 
Respondents' preference (% respondents)	29.8	70.2
Used in checklist	A	B

9.2.2 Ease of completion of each check item

The results presented in this section investigate whether there was a significant difference in ratings for ease of completion for each check item of checklist A compared to checklist B. Each checklist comprised 13 check items. The list below presents the topic areas of each of these. The actual descriptors of each check item used in the two checklists differ in terms of layout and information provided (e.g. numerical or word descriptor and illustration). For full versions of each check item used in the checklists, see Appendix 4.

- 1 Frequency – shoulder/arm movements
- 2 Repetition
- 3 Force
- 4 Posture – awkward head/neck posture
- 5 Posture – awkward back posture
- 6 Posture – awkward shoulder/arm posture
- 7 Posture – static shoulder and elbows
- 8 Posture – awkward wrist posture
- 9 posture – awkward hand/finger grip
- 10 Posture – static fingers, hand and wrist
- 11 Breaks
- 12 Work pace
- 13 Additional factors

The following results are from data collected from 88 completed assessments using checklist A and 86 completed assessments using checklist B. A Mann Whitney between-subjects statistical test was applied to the data. Results show that there were no significant differences in ratings for Checklist A and B for all check items except 4 and 5. These were rated as significantly more easy to complete on checklist B than checklist A.

When the data was split in relation to participant job position, it was found that there was no significant difference between the ease of completion ratings for each check item of Checklist A compared to Checklist B when completed by line leaders or line workers.

For checklist A on its own there was no significant difference between line leaders' and line workers' ratings of each check item, except for items 4 and 6, where line workers gave significantly higher (ie easier) ratings than line leaders.

For checklist B on its own there was no significant difference between line leaders' and line workers' ratings of each check item.

Level 2 analysis

The separate results for ease of each check item for each task are available from the authors on request.

9.2.3 Percentage agreement for each check item

This section investigates the level of agreement between participants' responses and the model response as to whether the risk factor was present or not present. This was calculated for each of the 13 check items. In total, data were collected from 90 completed assessments using checklist A and 86 for checklist B. Because the number of completed assessments using each checklist differed for each task (Table 9.17) the percentage values were calculated to enable direct comparison of the performance of each checklist for each of the 13 check items.

The results have been presented to enable comparisons of the level of agreement and the effect both of comparing the two checklists and of the participants' job position.

Comparing the checklists – analysis of all task data combined

This section compares the percentage of agreement between participants and the model response achieved for each checklist when the results from all four tasks are combined.

Table 9.18 shows that the level of agreement across the check items ranged from 47 to 99 per cent for checklist A and 38 to 95 per cent for checklist B. The responses of more than 80 per cent of participants matched the model response for both checklists for check items 1 and 11–13. It is important to note that these check items refer to aspects where there is a definitive answer. The other items, for which participants' responses agreed significantly less with the model response, tend to

Task	Worker group	Checklist A	Checklist B
Task 1	Team leaders, line leaders and line managers	10	8
	Line workers	24	16
	Unknown	1	0
Task 2	Team leaders, line leaders and line managers	5	5
	Line workers	13	6
	Unknown	0	0
Task 3	Team leaders, line leaders and line managers	7	10
	Line workers	17	24
	Unknown	0	0
Task 4	Team leaders, line leaders and line managers	2	5
	Line workers	11	12
	Unknown	0	0
TOTAL		90	86

Table 9.17
Number of completed assessments

Check item	Percentage agreement of participant responses to model responses (all tasks combined)	
	Checklist A (all companies, <i>n</i> = 90)	Checklist B (all companies, <i>n</i> = 86)
1	96	85
2	62	49
3	72	66
4	62	67
5	50	38
6	74	64
7	67	55
8	47	50
9	81	49
10	57	78
11	86	84
12	84	90
13	99	95

Table 9.18
Percentage agreements between participants and model responses for each check item for both checklists. Poor performance is shown in bold and major differences are tinted grey

require a judgment regarding body angles and ranges of motion rather than just time-based observation.

These poorer performing check items included 3, 4, 6 and 7, where 55 to 72 per cent of completed checklists agreed with the model response. Check items 2, 5 and 8, which are highlighted in Table 9.18, were particularly poor matches to the model responses, with fewer than half of the respondents providing matching responses.

It is interesting to note that for most check items, the performance of checklists A and B were very similar, except for items 9 and 10. Checklist A (81 per cent correct) significantly outperformed checklist B (49 per cent correct) for item 9. These results are shown in bold in Table 9.18. Figure 9.2 shows check item 9 from checklist A and B.

In contrast, Checklist B performed significantly better for item 10 (78 compared to 57 per cent correct). These results are also shown in bold in Table 9.18 and the treatments of this item in each checklist are shown in Figure 9.3.

Figure 9.2
Check item 9 from
the two checklists

9. Awkward hand / finger grip		
Both hands are not used to grip anything or they are using a 'Power grip' (Power grip is where the fingers are wrapped around an object and the thumb placed against it; used, for example, in certain hammering operations)		G 0
Is one of both hands using a 'Pinch' or 'Wide finger' grip for a part of the time (‘Pinch’ or ‘Wide finger’ grip is where the fingers are on one side of an object, and the thumb is on the other. Typically, an object lifted in a pinch or wide finger grip do not touch the palm)		A 1
Is one of both hands using a 'Pinch' or 'Wide finger' grip for more than half of the time (‘Pinch’ or ‘Wide finger’ grip is where the fingers are on one side of an object, and the thumb is on the other. Typically, an object lifted in a pinch or wide finger grip does not touch the palm)		R 2

Checklist A

9. Is a pinch or wide finger grip being used for more than 15% of the time	 Pinch  Wide finger	☐	☐
--	---	--------------------------	--------------------------

Checklist B

The effect of participants' job position – analysis of all task data combined

This section compares the percentage of agreement between participants and the model response achieved by each checklist by job position. In particular, it looks at the performance of leaders compared to workers when the results from all four tasks are combined. The results, presented in Table 9.19, show that overall the leaders' responses had a slightly higher level of agreement with the model responses for all check items when using checklist B compared to checklist A. Conversely, the workers gained a slightly higher level of agreement when using checklist A.

For each check item in each checklist there was only a small difference between the levels of agreement achieved by the leaders and the workers, except in the case of items 7 and 9 of checklist A and items 6 and 9 for checklist B. These are highlighted in Table 9.19. More workers (77 per cent) than leaders (50 per cent) completed item 9 on checklist A correctly, whereas for checklist B the results for the same item were reversed (71 per cent for leaders, 45 per cent for workers).

10. Static fingers, hand and wrist	
Both hands and wrists regularly change position to adopt relaxed/neutral postures during every hour of work	G 0
One or both hands and wrists are in a static position (i.e. infrequently moved) for between 1 and 2 consecutive hours	A 1
One or both hands and wrists are in a static position (i.e. infrequently moved) for more than 2 consecutive hours	R 2

Checklist A

10. Is one or both hands and wrists held in a static position (i.e. infrequently moved) for more than 1 hour		☐	☐
--	--	--------------------------	--------------------------

Checklist B

Figure 9.3
Check item 10
from the two
checklists

Check item	Percentage agreement of participant responses and model response for tasks 1 to 4			
	Checklist A		Checklist B	
	Team leader, line leader or line manager (n = 24)	Line worker or operative (n = 65)	Team leader, line leader or line manager (n = 28)	Line worker or operative (n = 58)
1	96	95	89	83
2	67	60	46	50
3	75	77	82	59
4	58	65	61	71
5	54	49	46	35
6	75	74	82	55
7	54	71	61	52
8	42	48	57	47
9	50	77	71	45
10	67	52	86	74
11	88	86	82	85
12	75	89	86	91
13	100	99	89	98

Table 9.19
Percentage agreements of participants with the model responses for each check item in each checklist, split by respondents' job position (key differences are shown in bold)

The effect of type of task assessed - analysis of separate task data

The results show that while there are differences in the level of agreement across tasks, with some tasks gaining fewer correct responses for some check items than others, overall similar patterns are seen across both checklists A and B. Full results are available from the authors on request.

9.2.4 Total number of risk factors

The following results are from data collected from 90 completed assessments using checklist A and 85 using checklist B. Table 9.20 presents descriptive statistics for the ADVs of the two checklists. Twenty-three per cent of all completed Checklist As agreed with the number of total risk factors present in the model response. The majority (54 per cent) provided a total number of risk factors present that varied by 1 or 2 from the model response. The remaining 22 per cent had discrepancies values of between 3 and 8. A similar level of performance was gained from Checklist B but with slightly fewer participants agreeing exactly with the model response (13 per cent of completed checklist Bs) and 58 per cent varying from the model response by 1 or 2 risk factors.

Table 9.20
Descriptive statistics for the ADVs for checklists A and B

	Absolute discrepancy values	
	Checklist A	Checklist B
Number of completed checklists	90	85
Mean	1.7	1.9
Median	1	2
Mode	1	1
Standard deviation	1.6	1.5
Minimum	0	0
Maximum	7	8

When split by job position, similar ranges of discrepancy values were found for leaders and workers (Table 9.21). For both groups, checklist A yielded a slightly higher percentage of correct responses (ie discrepancy value of 0), with 23 per cent for the leader group and 24 per cent for the workers. Checklist B yielded fewer correct values, with 7 per cent for the leaders and 16 per cent for the workers.

Table 9.21
Descriptive statistics for the ADVs for checklists A and B, split by job position

	Absolute discrepancy values			
	Checklist A		Checklist B	
	Team leader, line leader or line manager	Line worker or operative	Team leader, line leader or line manager	Line worker or operative
Number of completed checklists	26	57	28	57
Mean	1.8	1.7	2.1	1.8
Median	1	2	2	1
Mode	1	2	2	1
Standard deviation	1.8	1.5	1.6	1.5
Minimum	0	0	0	0
Maximum	6	7	8	7

To investigate whether the checklists differed in terms of a tendency to either over- or under-assess the total number of risk factors present, the negative values of the discrepancy values were also considered (Table 9.22).

	Discrepancy values	
	Checklist A	Checklist B
Number of completed checklists	90	85
Mean	0.9	-0.14
Median	1	0
Mode	0	1
Standard deviation	2.2	2.5
Minimum	-4	-8
Maximum	7	5

Table 9.22
Descriptive statistics for the discrepancy values for checklists A and B

A negative discrepancy value means that the participant under-assessed the total number of risk factors present, whereas a positive value means that the participant over-assessed them. The results show that 45 per cent of the checklist B users tended to under-assess the total number of risk factors present. In comparison, only 23 per cent of Checklist A users did this.

Table 9.23 shows the discrepancy values when split by job position. Here it was found that for checklist A a slightly higher percentage of the leaders tended to under-assess the number of risk factors present, with 35 per cent gaining discrepancy values between -1 and -4. Only 19 per cent of the workers under-assessed the number of risk factors present (-1 to -4). Conversely, the workers tended to over-assess the number of risk factors present more than the leaders (57 vs 42 per cent).

The figures for Checklist B show that a similar percentage of both groups under- and over-assessed the total number of risk factors present.

	Discrepancy values			
	Checklist A		Checklist B	
	Team leader, line leader or line manager	Line worker or operative	Team leader, line leader or line manager	Line worker or operative
Number of completed checklists	26	63	28	57
Mean	0.9	0.9	-0.3	-0.1
Median	0	1	-0.5	
Mode	-1	0	-1	1
Std. Deviation	2.4	2.1	2.7	2.4
Minimum	-3	-4	-8	-7
Maximum	6	7	4	5

Table 9.23
Descriptive statistics for the discrepancy values for checklists A and B, split by job position

Statistical analysis of the ADVs

Statistical analysis of the ADVs supports the aforementioned percentage results. It was found that there was no statistically significant difference between the checklists in terms of agreement with the model response for the total number of risk factors present. It also confirmed there was no statistically significant differences between the two job position groups on the same parameter (Table 9.24). Both checklist and job position groups performed equally well.

Table 9.24
Results from the
statistical analysis
of the ADVs

Effect or interaction investigated	Question to answer	Statistical result
Main effect of checklist on ADVs	On average, before training, in what way do the ADVs change with checklist, regardless of the effect (if any) of other variables?	No significant difference between checklists ($p = 0.166$)
Main effect of company on ADVs	On average, before training, in what way do the ADVs change with company, regardless of the effect (if any) of other variables?	No significant difference between companies ($p = 0.105$)
Main effect of job position on ADVs	On average, before training, in what way do the ADVs change with job position, regardless of the effect (if any) of other variables?	No significant difference between job position groups (team leaders, line leaders, line managers group and line workers group) ($p = 0.698$)
Estimated interaction of company and job position on ADVs	On average, before training, to what extent is the effect of company on the ADVs modified by the effect of job position, regardless of the effect of other variables?	No significant effect ($p = 0.280$)
Estimated interaction of job position and checklist on ADVs	On average, before training, to what extent is the effect of job position on the ADVs modified by the effect of checklist, regardless of the effect of other variables?	No significant effect ($p = 0.136$)

Level 2 – Analysis of data for each task separately

When comparing the data by job position for each task separately, it was possible to identify a significant difference between checklists A and B. It can be seen in Table 9.25 that Checklist B generated significantly lower discrepancy values for task 3 when completed by the workers than the leaders. This suggests that checklist B offers a greater opportunity for workers to obtain scores for task 3 closer to those identified as correct by the expert assessors.

This was the only statistically significant finding for this section of the Level 2 analysis, although the remaining results are available from the authors on request.

Table 9.25
Statistical results
for task 3,
comparing
individuals by job
position

	Team leader, line leader or line manager ($n = 8$)	Line worker or operative ($n = 15$)
Mean	2.3	1.25
Standard deviation	2.11	1.29
Minimum	1	0
Maximum	8	6
Overall significance in Checklist B results	0.055	

9.2.5 Overall risk ratings

The following results are from data collected from 90 completed assessments using Checklist A and 86 using checklist B. Table 9.26 shows that 71 per cent of checklist A answers agreed with the model response in terms of overall risk rating, whereas only 38 per cent of completed checklist Bs agreed with the model response. The majority of the remaining completed checklists (both A and B) over-rated the level of risk of the assessed tasks.

	Percentage of completed assessments	
Percentage of completed assessments that:	Assessment A (n = 90)	Assessment B (n = 86)
Agreed with model response	71	38
Under rated the risk	2	12
Over rated the risk	26	34
No risk rating given	1	16

Table 9.26
Percentage of completed checklists with risk ratings and how they related to the model response

Comparison of the assessment results for each checklist for overall risk rating when split by job position showed that there was little difference between the job position groups, with similar percentages of completed assessments by both job groups agreeing with the model response for both checklists. This is shown in Table 9.27.

	Percentage of respondents			
	Assessment A		Assessment B	
	Team leader, line leader or line manager (n = 24)	Line worker or operative (n = 65)	Team leader, line leader or line manager (n = 28)	Line worker or operative (n = 58)
Agreed with model response	63	74	32	41
Underrated the risk	4	2	3	16
Overrated the risk	33	23	36	33
No risk rating given	0	2	29	11

Table 9.27
Percentage of completed checklists with risk ratings and how they related to the model response when split by job position

9.2.6 Suggestions for improvement

Table 9.28 shows that checklist B produced a greater number of suggestions for improvement than checklist A. However, when these suggestions were further developed by providing more in-depth descriptions of the types of change to be made, both checklists performed equally well, with 41 per cent of completed Checklist As providing more in-depth suggestions and 42 per cent of checklist Bs.

Table 9.29 presents the descriptive statistics for the number of suggested changes to task activities which are intended to reduce the level of MSD risks. These are split by job position and checklist. It can be seen that checklist B encouraged a significantly larger number of suggestions from both worker types. Table 9.30 shows the percentages of respondents who provided additional suggestions from their own initiative or provided more in-depth descriptions of changes. Here there is little difference between the two checklists, although it should be noted that the percentage of leaders making such comments was approximately half that of workers.

Table 9.28

Descriptive statistics for the number of suggested changes to reduce risks when using each checklist

	Number of suggested changes to reduce the risks	
	Checklist A	Checklist B
Number of completed checklists	90	86
Mean	0.96	9.24
Median	0.00	8.00
Mode	0	10
Standard deviation	1.373	6.608
Minimum	0	0
Maximum	6	25

Table 9.29

Descriptive statistics for the number of suggested changes to reduce risks when using each checklist, split by job position

	Number of suggested changes to reduce the risks			
	Checklist A		Checklist B	
	Team leader, line leader or line manager	Line worker or operative	Team leader, line leader or line manager	Line worker or operative
Number of completed checklists	24	65	28	58
Mean	0.58	1.1	7.11	10.28
Median	0	0	6.50	9.50
Mode	0	0	0	10
Standard deviation	1.1	1.5	5.315	6.958
Minimum	0	0	0	0
Maximum	3	6	21	25

Table 9.30

Percentage of respondents who made their own suggestions or wrote more in-depth descriptions of possible changes

Checklist	All participants		Team leader, line leader or line manager		Line worker	
	A	B	A	B	A	B
Percentage of participants	41	42	25	29	48	48

9.3 Trial 2 – Comparing trained and untrained users

The following results are from data collected from a total of 26 participants and 46 completed checklists. Table 9.31 shows the number of completed checklists included in the analysis for tasks 1 and 2.

	Number of completed A and B assessments			
	Task 1		Task 2	
	Before training	After training	Before training	After training
Checklist A	12	12	12	16
Checklist B	10	10	8	10

Table 9.31
Ratings for ease of completing each check item of checklists A and B when assessing task 1 before and after training

The discrepancy between the number of participants and the number of checklists is accounted for by staff turnover between evaluations (trained and untrained). Since most of the participating companies employed a large number of temporary workers and are engaged in highly seasonal trades, the changes in staff profile are more significant than in some other market segments.

9.3.1 Ease of completion of each check item

For convenience, the check items that appear on both checklists are repeated below:

- 1 Frequency – shoulder/arm movements
- 2 Repetition
- 3 Force
- 4 Posture – awkward head/neck posture
- 5 Posture – awkward back posture
- 6 Posture – awkward shoulder/arm posture
- 7 Posture – static shoulder and elbows
- 8 Posture – awkward wrist posture
- 9 posture – awkward hand/finger grip
- 10 Posture – static fingers, hand and wrist
- 11 Breaks
- 12 Work pace
- 13 Additional factors

The ease of completion of the two checklists was compared for the two chosen activities (tasks 1 and 2). This was compared between trained and untrained participants to search for significant differences. Tables 9.32 and 9.34 present the results for checklist A and Tables 9.33 and 9.35 present those for checklist B.

Only two areas of significant difference were observed between the groups. For task 1 a difference was noted for check item 3. Here checklist B revealed a higher typical score with training. For task 2 a difference was observed for check item 9. In this instance training resulted in a lower score being recorded for checklist A.

9.3.2 Percentage agreement for each check item

A comparison was made between the participants' performance against the model responses recorded by the expert panel.

Task 1

Table 9.36 shows the expert responses for Task 1. Table 9.37 shows that the level of participant agreement with the model responses across check items ranged from 33 to 92 per cent for checklist A before training and from 50 to 100 per cent after training. There was substantial improvement after training for check items 2, 5 and 11, and substantially worsened performance following training for check items 3 and 8. Training did not appear to have a notable effect for check items which previously gained high performance (ie those which achieved 80 per cent or more agreement before training). This included check items 1, 4, 6, 8 and 9. Training also did not appear to have an effect on the poorer performing check items, 7 and 10 (which yielded less than 70 per cent agreement before training).

Table 9.32
Ratings for ease of completing each check item of checklist A when assessing task 1 before and after training

	Task 1 – Checklist A: Ratings for ease of completing each check item										
	Before training					After training					Significant?
	Mean	Mode	SD	Min	Max	Mean	Mode	SD	Min	Max	
1	3.75	3	0.75	3	5	3.67	4	0.89	2	5	No
2	3.75	3	1.06	2	5	3.67	4	1.37	1	5	No
3	4.08	4	0.79	3	5	3.91	4	0.94	2	5	No
4	3.25	3	0.62	2	4	3.17	3	1.03	1	5	No
5	3.83	4	0.94	2	5	3.83	3	0.83	3	5	No
6	3.67	5	1.23	2	5	3.67	4	0.89	2	5	No
7	3.33	3	1.15	2	5	2.83	3	1.19	1	5	No
8	3.50	3	0.90	2	5	3.27	3	0.90	2	5	No
9	3.67	4	1.15	2	5	3.50	3	0.90	2	5	No
10	3.58	5	1.24	2	5	3.17	2	1.19	2	5	No
11	4.00	5	1.13	2	5	4.42	5	0.67	3	5	No
12	3.92	5	1.51	1	5	4.25	5	1.22	1	5	No
13	3.58	5	1.38	2	5	3.58	4	1.00	2	5	No

Table 9.33
Ratings for ease of completing each check item of checklist B when assessing task 1 before and after training

	Task 1 – Checklist B: Ratings for ease of completing each check item										
	Before training					After training					Significant?
	Mean	Mode	SD	Min	Max	Mean	Mode	SD	Min	Max	
1	3.50	3	1.35	1	5	4.1	4	0.74	3	5	No
2	3.40	2	1.27	2	5	4.0	4	0.82	3	5	No
3	3.50	3	0.71	3	5	4.0	4	0.67	3	5	$p = 0.096$
4	3.60	4	0.52	3	4	3.6	4	0.84	2	5	No
5	3.67	4	0.87	2	5	3.7	4	1.16	2	5	No
6	3.00	4	1.12	1	4	3.6	3	0.97	2	5	No
7	3.25	4	0.89	2	4	3.6	3	0.70	3	5	No
8	3.00	3	1.00	2	5	3.2	2	1.32	2	5	No
9	3.22	3	0.83	2	5	3.8	3	1.14	2	5	No
10	3.50	3	0.53	3	4	3.7	3	0.82	3	5	No
11	4.40	4	0.52	4	5	3.7	3	0.95	3	5	No
12	4.11	4	1.27	1	5	4.0	4	0.82	3	5	No
13	3.30	3	1.16	1	5	3.6	3	0.84	3	5	No

	Task 2 – Checklist A: Ratings for ease of completing each check item										
	Before training					After training					Significant?
	Mean	Mode	SD	Min	Max	Mean	Mode	SD	Min	Max	
1	4.00	4	0.74	3	5	3.67	3	0.82	3	5	No
2	3.42	4	0.90	2	5	3.57	4	0.53	3	4	No
3	4.00	4	0.60	3	5	3.43	4	1.27	1	5	No
4	3.58	4	1.00	2	5	4.14	5	0.90	3	5	No
5	3.33	4	0.78	2	4	4.00	4	0.82	3	5	No
6	3.58	4	0.79	2	5	4.14	4	0.38	4	5	No
7	3.42	4	0.79	2	4	3.86	3	0.90	3	5	No
8	3.58	4	0.51	3	4	3.71	3, 4	0.76	3	5	No
9	3.58	4	0.90	2	5	2.29	1, 2, 3	1.11	1	4	$p = 0.022$
10	3.55	4	0.82	2	5	3.14	3	1.07	1	4	No
11	4.25	5	0.97	2	5	4.50	5	0.84	3	5	No
12	4.25	5	0.97	2	5	4.67	5	0.52	4	5	No
13	3.67	4	0.89	2	5	4.00	5	0.82	3	5	No

Table 9.34
Ratings for ease of completing each check item of checklist A when assessing task 2 before and after training

	Task 2 – Checklist B: Ratings for ease of completing each check item										
	Before training					After training					Significant?
	Mean	Mode	SD	Min	Max	Mean	Mode	SD	Min	Max	
1	4.25	5	1.17	2	5	4.22	5	0.83	3	5	No
2	4.25	5	1.17	2	5	4.11	5	0.93	3	5	No
3	4.38	5	0.92	3	5	4.11	5	0.93	3	5	No
4	4.13	5	0.99	3	5	3.56	3	1.01	2	5	No
5	4.13	5	1.13	2	5	3.89	4	.78	3	5	No
6	3.88	3	0.99	3	5	3.56	3	1.24	2	5	No
7	4.25	5	0.89	3	5	3.67	3	1.12	2	5	No
8	2.75	1	1.58	1	5	3.22	3	1.48	1	5	No
9	4.00	4	1.07	2	5	3.89	4	0.78	3	5	No
10	3.75	4	1.04	2	5	3.22	2	1.39	1	5	No
11	4.38	5	0.74	3	5	4.00	3	1.00	3	5	No
12	4.38	5	0.74	3	5	3.67	5	1.23	2	5	No
13	3.5	3	0.93	2	5	3.44	3	1.01	2	5	No

Table 9.35
Ratings for ease of completing each check item of checklist B when assessing task 2 before and after training

Table 9.36
Model response for
task 1

Question number	Risk status
Q1	Present
Q2	Risk not present
Q3	Risk not present
Q4	Present
Q5	Risk not present
Q6	Present
Q7	Risk not present
Q8	Present
Q9	Present
Q10	Risk not present
Q11	Risk not present
Q12	Risk not present
Q13	Present

Table 9.38 extends this comparison for checklist B. Here, an improvement in performance was demonstrated after training for check items 8 and 9. However, worse performance was noted after training for check items 5, 7 and 10. These changes were noted in the leader group but not replicated by the workers.

Table 9.37
Percentage
agreement of
participants to the
model response for
task 1 using
checklist A, split by
job position.
Significant
improvements are
shaded dark grey
and significant
reductions in light
grey

Check item	All participants (n = 12)		Team leader, line leader or line manager (n = 5)		Line worker or operative (n = 7)	
	Before training	After training	Before training	After training	Before training	After training
1	92	92	80	100	100	86
2	33	50	20	40	43	57
3	83	58	80	40.0	86	71
4	92	83	100	80	86	86
5	33	67	40	60	29	71
6	92	100	100	100	86	100
7	67	67	40	60	86	71
8	83	75	60	80	100	71
9	83	83	60	60	100	100
10	58	58	60	60	57	57
11	67	92	40	80	86	100
12	75	83	40	60	100	100

Check item	All participants (n = 10)		Team leader, line leader or line manager (n = 4)		Line worker or operative (n = 6)	
	Before training	After training	Before training	After training	Before training	After training
1	80	90	100	100	67	83
2	10	0	0	0	17	0
3	50	30	50	25	50	33
4	80	90	75	100	83	83
5	70	30	100	25	50	33
6	80	90	100	100	67	83
7	80	30	100	25	67	33
8	40	70	50	100	33	50
9	70	90	75	100	67	83
10	70	20	100	0	50	33
11	80	100	75	100	83	100
12	90	90	100	100	83	83

Table 9.38
Percentage agreement of participants to the model response for task 1 using checklist B, split by job position. Significant improvements are shaded dark grey and significant reductions in light grey

Task 2

The exercise was repeated for task 2. Table 9.39 shows the model response for task 2 against which participant performance was evaluated.

Question number	Risk status
Q1	Present
Q2	present
Q3	Risk not present
Q4	Present
Q5	present
Q6	Present
Q7	present
Q8	Risk not present
Q9	Present
Q10	Present
Q11	Present
Q12	Present

Table 9.39
Model response for task 2

For task 2, a different pattern of performance was observed. Table 9.40 shows that for checklist A, leaders demonstrated improved performance after training for check items 4 and 9. However, decreased post-training performance for both leaders and workers was observed for check items 3, 5 and 12. The workers additionally recorded worse post-training performance for check items 4 and 6.

Table 9.40
Percentage agreement of participants to the model response for task 2 using checklist A, split by job position. Significant improvements are shaded dark grey and significant reductions in light grey

Check item	All participants		Team leader, line leader or line manager		Line worker or operative	
	Before training	After training	Before training	After training	Before training	After training
1	100	100	100	100	100	100
2	100	94	100	80	100	100
3	83	69	100	60	80	63
4	58	50	0	40	70	38
5	75	56	100	60	70	38
6	92	69	100	80	90	63
7	58	63	50	60	60	88
8	42	38	50	40	40	50
9	83	88	50	100	90	100
10	50	81	100	80	40	88
11	92	100	100	100	90	100
12	100	50	100	60	100	25

Table 9.41 goes on to show the results for checklist B. Here it can be seen that the leaders increased performance with training for check items 4 and 9. The workers saw a fall in performance for check item 4, but with training improved on item 10. Leaders and workers both showed a decrease in post-training performance for items 5 and 12.

9.3.3 Total number of risk factors

The following results are from data collected from 24 completed assessments using checklist A and 20 completed assessments using checklist B. The data are presented in two forms, each looking at the effect of training: firstly, an evaluation of overall absolute discrepancy values split by checklist and by job position, and secondly, the discrepancy values expressed as positive or negative errors (under- or over-reporting of apparent hazards) are presented, again by checklist and job position. The data are presented as both tables and histograms to give a clearer indication of the performance achieved.

Task 1

Checklist A – absolute discrepancy values

Table 9.42 presents descriptive statistics for the discrepancy values (participant total number of risk factors present minus model response total number of risk factors present) of checklists A and B for task 1. Figure 9.4 presents a histogram depicting the percentage of completed checklist and the range of discrepancy values gained from each.

It can be seen that 42 per cent of those completing checklist A were 3 or more away from the model response before training. After training this improved and was reduced to 25 per cent. Seventeen per cent of all who completed checklist A before training agreed with the number of total risk factors present in the model response (ie they had a discrepancy value of 0) whereas after training this increased to 33 per cent.

Check item	All participants		Team leader, line leader or line manager		Line worker or operative	
	Before training	After training	Before training	After training	Before training	After training
1	100	89	100	100	100	100
2	100	89	100	80	100	100
3	88	56	100	60	80	63
4	75	89	0	40	70	38
5	88	100	100	60	70	38
6	63	89	100	80	90	63
7	38	44	50	60	60	88
8	63	56	50	40	40	50
9	50	89	50	100	90	100
10	38	78	100	80	40	88
11	100	89	100	100	90	100
12	88	33	100	60	100	25

Table 9.41
Percentage agreement of participants to the model response for task 2 using checklist B, split by job position. Significant improvements are shaded dark grey and significant reductions in light grey

	Absolute discrepancy values			
	Checklist A		Checklist B	
	Before training	After training	Before training	After training
Number of completed checklists	12	12	10	10
Mean	2.5	1.8	2	4.9
Median	2	1	2	4.5
Mode	1	0 and 1	2	4
Std. Deviation	2.1	2.2	1.15	2.69
Minimum	0	0	0	0
Maximum	6	7	4	9

Table 9.42
Descriptive statistics for each checklist's ADVs for task 1

For Checklist B the performance is less positive. While 30 per cent of participants had discrepancy scores of 3 or more before training, this rose to 90 per cent after training. This suggests that reporting error was encouraged by the training – most likely through raising awareness such that false positives are recorded.

Tables 9.43 and 9.44 present the descriptive statistics split by job position for Task 1 when using checklist A and checklist B. Figure 9.5 presents the performance as percentages when split by worker type. Here it can be seen that for checklist A the line workers achieved the greatest improvement in accurately identifying the correct number of risk factors.

For checklist B, line leaders showed a large increase in discrepancy values after training, moving from 0 per cent recording values of 3 or more before training to 100 per cent afterwards. Line workers were split in their training results, with correct values moving from 0 to 17 per cent, but also values above 3 moving from 50 to 84 per cent.

Figure 9.4
Percentage of participants' ADVs for each checklist when used to assess task 1

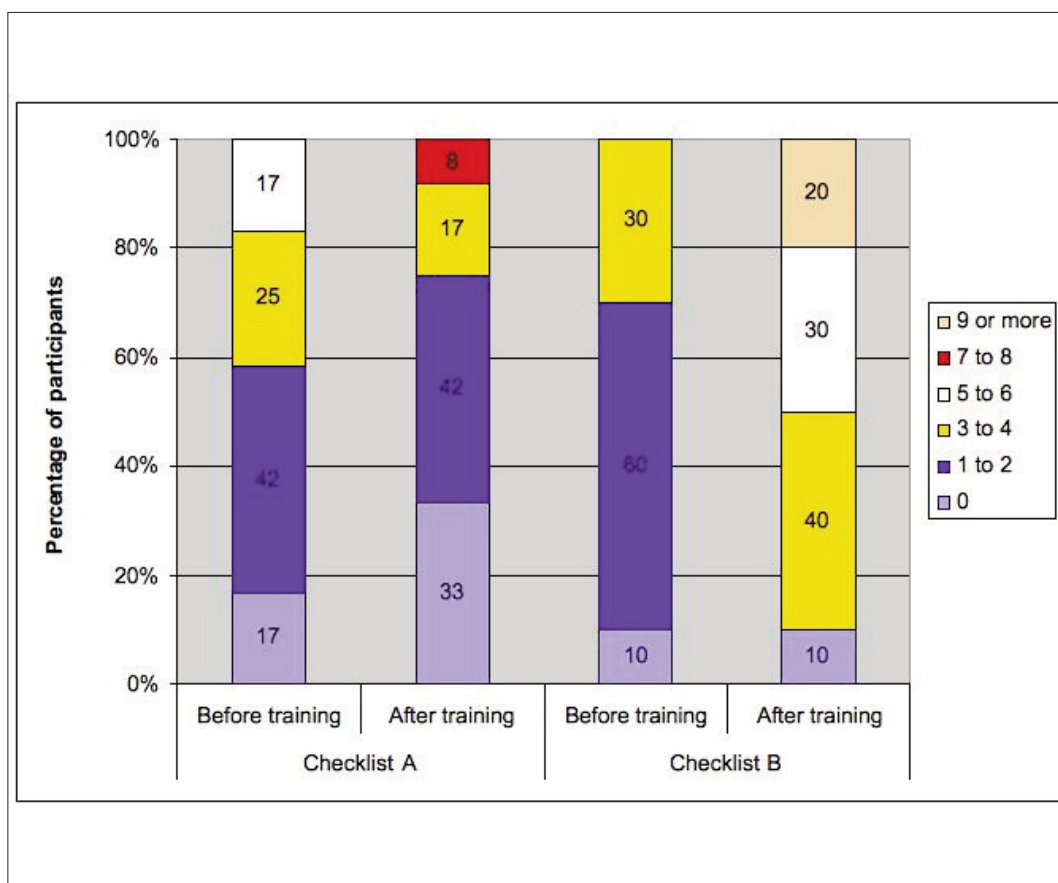


Table 9.43
Descriptive statistics for checklist A ADVs, task 1

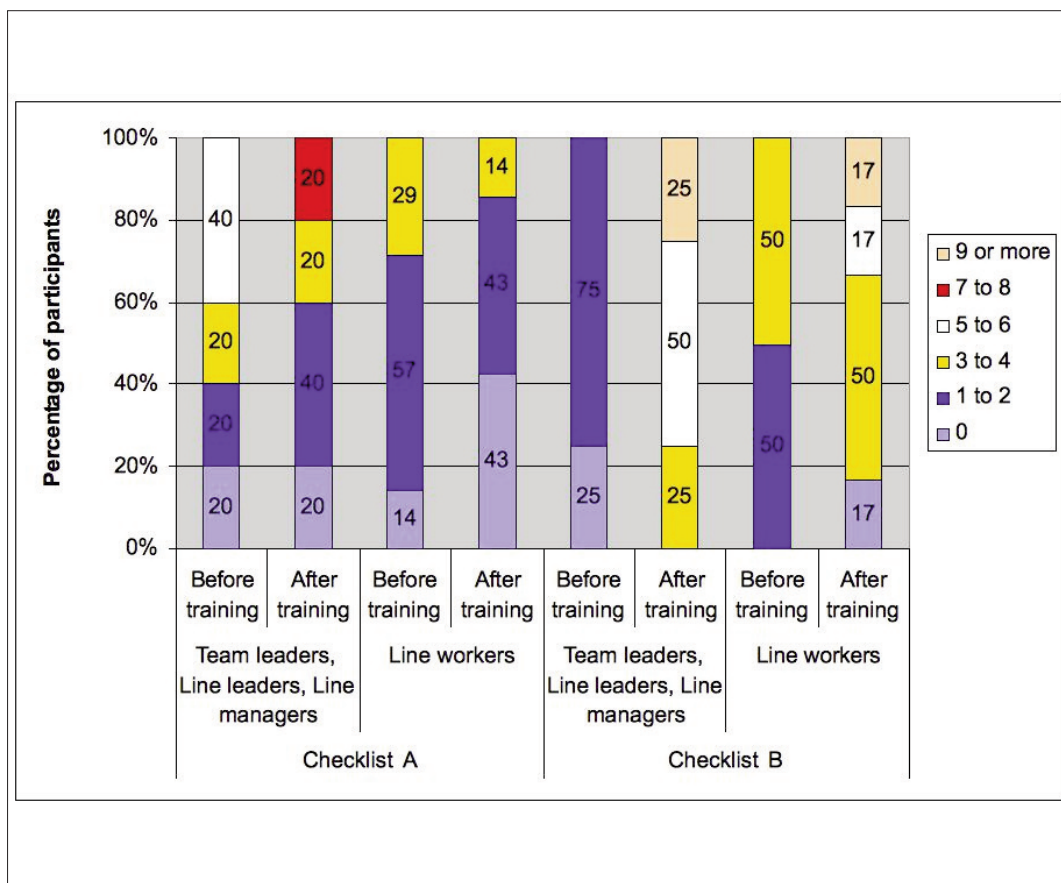
	Absolute discrepancy values			
	Team leader, line leader, line manager		Line worker, line operative	
	Before training	After training	Before training	After training
Number of participants	5	5	7	7
Mean	3.4	2.6	1.8	1.1
Median	4	1	2	1
Mode	6	1	1 and 2	0
Standard deviation	2.8	2.9	1.4	1.5
Minimum	0	0	0	0
Maximum	6	7	4	4

The performance was further assessed to establish whether the changes in performance were attributable to over- or under-reporting of apparent risk factors. This was achieved by breaking the discrepancy values into positive and negative bands, representing the drift of participant scores from the expert scores in the model assessment. Table 9.45 gives the descriptive statistics for these revised discrepancy values for checklists A and B before and after training.

The positive and negative discrepancy values are shown in Figure 9.6. The green band illustrates where the participants agreed with the expert scores. It can be seen that initially for checklist A, 8 per cent of the responses were due to under-reporting and this continued after training. Also after training, while there was an increase in the correct score (17 to 33 per cent) there was also an

	Absolute discrepancy values			
	Team leader, line leader, line manager		Line worker, line operative	
	Before training	After training	Before training	After training
Number of participants	4	4	6	6
Mean	1.5	6	2.3	4.2
Median	2	5.5	2.5	4
Mode	2	4	1	4
Standard deviation	1	2.2	1.21	2.93
Minimum	0	4	1	0
Maximum	2	9	4	9

Table 9.44
Descriptive
statistics for
checklist B ADVs,
task 1



increase in the percentage of respondents over-reporting risks by 7 or more. This suggests that some participants were over-sensitised to risk factors by the training process.

For checklist B, under-reporting was eliminated by training (30 to 0 per cent). However, the accurate response rate did not change (10 per cent before and after) with over-reporting rising significantly. Again, it appears that training has artificially raised awareness of potential risks.

Tables 9.46 and 9.47 present the descriptive statistics for Task 1 positive and negative discrepancy values when split by job position. Figure 9.7 goes on to present the positive and negative discrepancy values for both checklists by job position.

Table 9.45
Positive and negative discrepancy values for each checklist, task 1

	Discrepancy values (positive and negative values)			
	Checklist A		Checklist B	
	Before training	After training	Before training	After training
Number of completed checklists	12	12	10	10
Mean	2.3	1.6	0.8	4.9
Median	2	1	1.5	4.5
Mode	0	0	2	4
Standard deviation	2.3	2.3	2.25	2.69
Minimum	-1	-1	-3	0
Maximum	6	7	4	9

Figure 9.6
Percentage of participants' positive and negative discrepancy values for each checklist when used to assess task 1

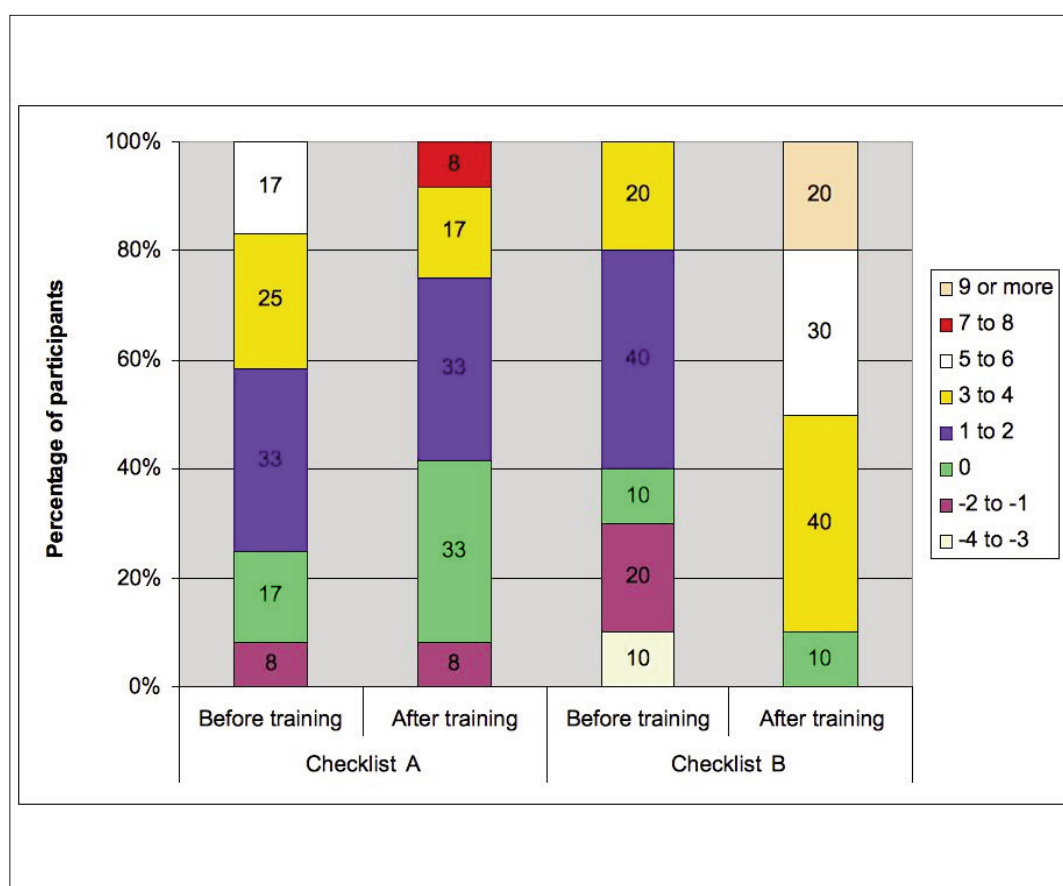


Figure 9.7 shows that for checklist A, line workers recorded no negative discrepancy values, and their increase in accurate responses following training is accounted for in a reduction in both high and moderate over-reporting. Line leaders continued to under-report after training (20 per cent), but the over-reporting distribution widened. This suggests that training resulted in a wider diversity of positive responses.

For checklist B, line workers moved from a split between under-reporting (34 per cent) and over-reporting (66 per cent) to a wider spread of responses after training. Once trained, 17 per cent of this group recorded scores that matched the expert score. However, over-reported errors moved from 33 per cent recording 3 or more to 84 per cent recording 3 or more. This is strong evidence for over-

	Discrepancy values (positive and negative values)			
	Team leader, line leader, line manager		Line worker, line operative	
	Before training	After training	Before training	After training
Number of participants	5	5	7	7
Mean	3	2.2	1.9	1.4
Median	4	1	2	1
Mode	6	-1	1	0
Standard deviation	3.3	3.3	1.4	1.5
Minimum	-1	-1	0	0
Maximum	6	7	4	4

Table 9.46
Positive and negative discrepancy values for checklist A, task 1, by job position

	Discrepancy values (positive and negative values)			
	Team leader, line leader, line manager		Line worker, line operative	
	Before training	After training	Before training	After training
Number of participants	4	4	6	6
Mean	0.5	6	1	4.2
Median	1	5.5	1.5	4
Mode	2	4	-3	4
Standard deviation	1.91	2.16	2.61	2.93
Minimum	-2	4	-3	0
Maximum	2	9	4	9

Table 9.47
Positive and negative discrepancy values for checklist B, task 1, by job position

sensitising during training. Line leaders using checklist B moved from a 100 per cent spread around the expert score (a discrepancy value of plus or minus 2) to 100 per cent recording positive discrepancy values of 3 or more. This suggests that line leaders were more receptive to over sensitisation by training.

Task 2

A similar analysis was undertaken of the results for Task 2. Again this considered absolute discrepancy values as well as positive and negative discrepancy values for checklists A and B. Each checklist was evaluated by job position both before and after training.

Task 2 (absolute discrepancy values)

Table 9.48 presents the descriptive statistics for the Discrepancy Values for Task 2 and the percentile results for the task 2 absolute discrepancy values are shown in Figure 9.8. Here it can be seen that initially, when using checklist A, participants recorded parity with the experts in 17 per cent of the evaluations. This reduced on training to 6 per cent, with a corresponding increase in absolute discrepancy values. It is notable that there is an increase in the spread of absolute discrepancy values moving from 25 per cent for 3–4 up to 6 per cent for 5–6. There is a corresponding increase in scores of 1–2 from 58 to 69 per cent. For checklist B, there was a similar increase in the upward drift of over representation of errors (0 to 25 per cent for values of 7 and more) but this is mitigated by an increase in matches to the expert scores, which went from 0 per cent without training to 25 per cent with training.

Figure 9.7
Percentage of participants' positive and negative discrepancy values for each checklist when used to assess task 1, split by job position

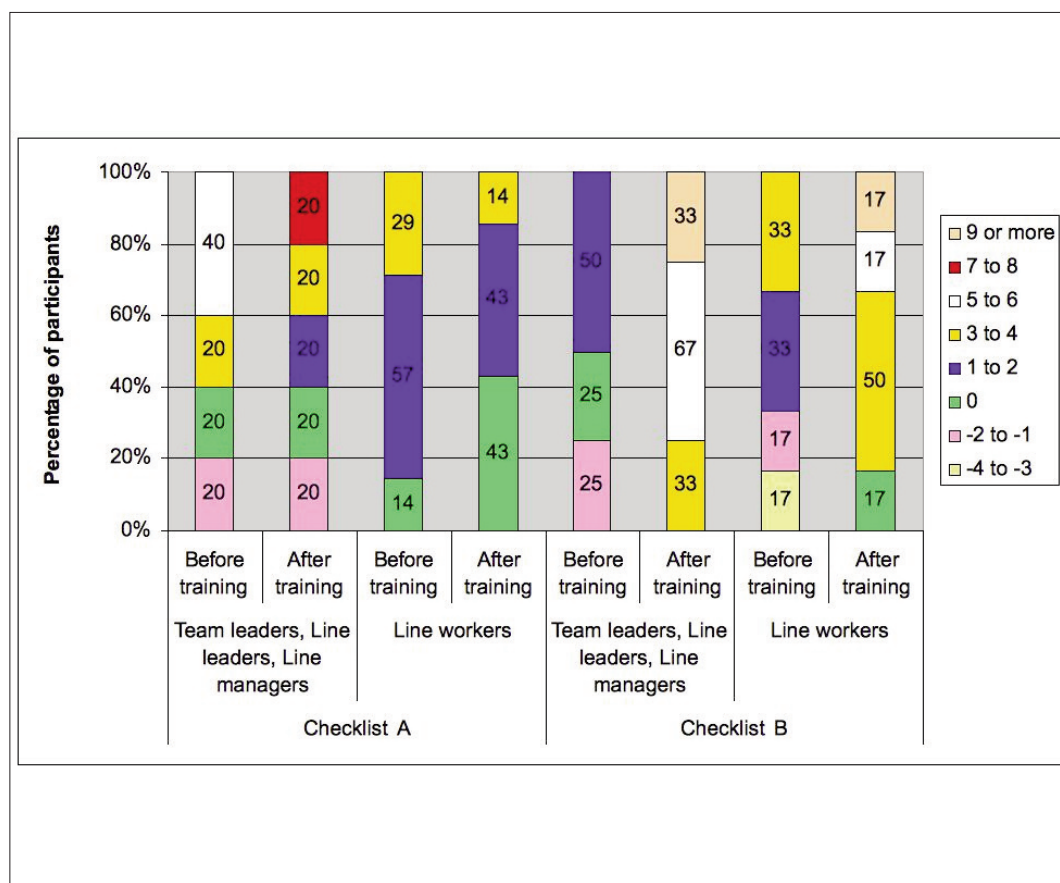


Table 9.48
Descriptive statistics for each checklist's ADVs for task 1

	Absolute discrepancy values			
	Checklist A		Checklist B	
	Before training	After training	Before training	After training
Number of completed checklists	12	16	8	8
Mean	2	2.1	2.25	2.4
Median	2	2	2.5	1
Mode	2	2	1	1
Standard deviation	1.4	1.2	1.16	3.2
Minimum	0	0	1	0
Maximum	4	5	4	8

Table 9.49 gives the descriptive statistics for checklist A in the Task 2 evaluation. The results are presented by job position. Table 9.50 offers the same information for checklist B.

These data are presented in graphical form in Figure 9.9. Here it can be seen that for checklist A there is a deleterious effect from training. For line leaders the drift was from 100 per cent recording only 1–2 errors to 20 per cent recording over 5 errors with the remainder still recording 1–2 errors. This is matched by the line workers. Initially, 20 per cent of this group recorded correct scores. However, with training, this dropped to zero, with 75 per cent scoring 1–2 errors and the number recording 3 or more remaining more or less constant. This further suggests that training has a negative effect on risk factor identification for this checklist.

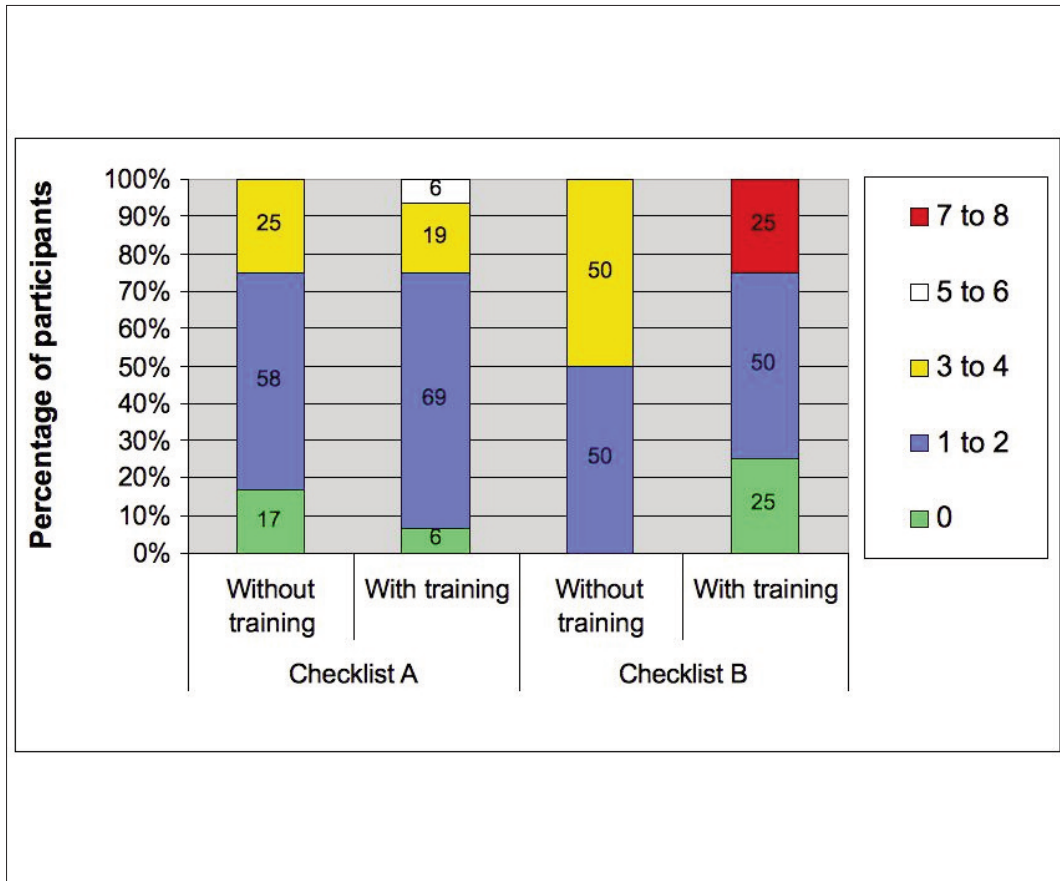


Figure 9.8
Percentage of participants' ADVs for each checklist when used to assess task 2

	Absolute discrepancy values			
	Team leader, line leader, line manager		Line worker, line operative	
	With training	Without training	Before training	Without training
Number of participants	2	5	10	8
Mean	1.5	2.2	2.1	2.25
Median	1.5	2	2	2
Mode	1	1	2	2
Standard deviation	0.71	1.64	1.52	0.89
Minimum	1	1	0	1
Maximum	2	5	4	4

Table 9.49
Checklist A statistics for task 2 by job position

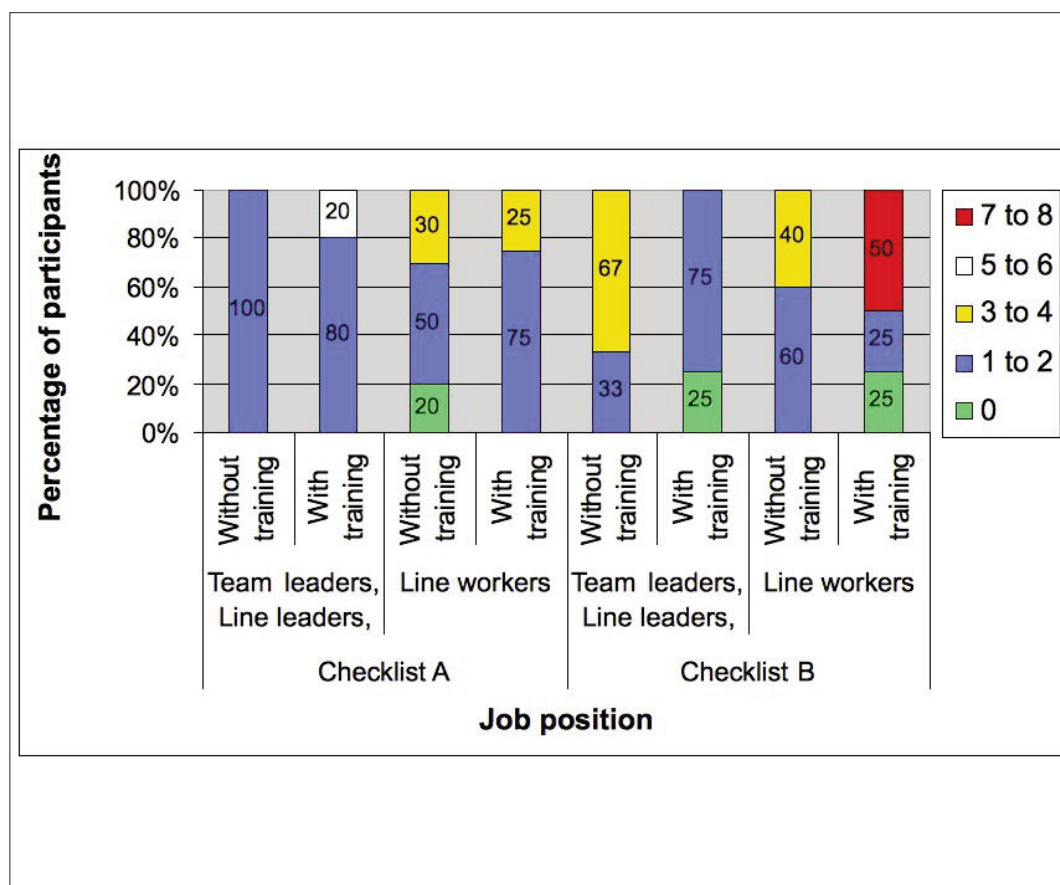
For checklist B there was a different trend in the number of errors after training. For line leaders there was reduction in the error rate, falling from 67 to 0 per cent recording 3 or more errors. It is also worthy of note that 25 per cent of the line leader participants were able to generate no errors following training whereas none had before.

Line workers saw a similar generation of correct scores following training (25 per cent) but this was offset by an increase in the number of errors for the remaining participants. 50 per cent of respondents recorded 7 or more errors after training, whereas the highest error rate prior to training had been 4 errors (40 per cent). Training again seems only partially effective here, with some users learning to expert level while others become more error prone.

Table 9.50
Checklist B
statistics for task 2
by job position

	Absolute discrepancy values			
	Team leader, line leader, line manager		Line worker, line operative	
	With training	Without training	Before training	Without training
Number of participants	3	4	5	4
Mean	-2.67	0.75	-2	0.5
Median	-3	1	-2	0.5
Mode	-4	1	-3	7
Standard deviation	1.53	0.5	1	6.14
Minimum	-4	0	-3	-7
Maximum	-1	1	-1	8

Figure 9.9
Percentage of
participants' ADVs
for each checklist
when used to
assess task 2, split
by job position

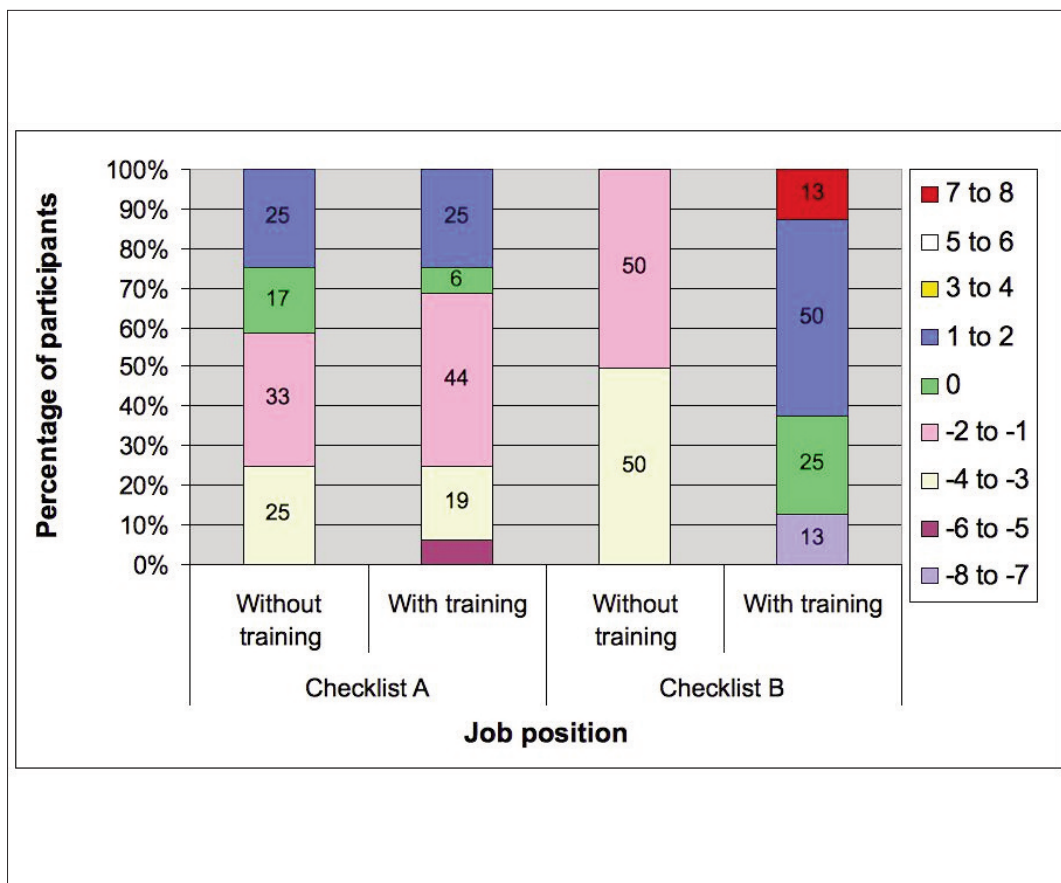


The data were then scrutinised to see whether the discrepancy values were either positive or negative – indicating whether participants were under-detecting risk factors or over-identifying these problems. Table 9.51 shows the descriptive statistics for these positive and negative discrepancy values.

The percentile values are shown in the histogram in Figure 9.10. Here it can be seen that checklist A generates significant under-reporting of risk factors with 58 per cent of respondents without training and 69 per cent of trained respondents identifying fewer risk factors than the expert score. The amount of over-representation remained the same at 25 per cent, suggesting that training caused participants to overlook potential health threats. Checklist B generated a much larger effect with

	Discrepancy values (positive and negative values)			
	Checklist A		Checklist B	
	With training	Without training	Before training	Without training
Number of completed checklists	12	16	8	8
Mean	-1.2	-1.3	-2.25	0.63
Median	-1.5	-2	-2.5	1
Mode	-4	-2	-3	1
Standard deviation	2.2	2.1	1.16	4
Minimum	-4	-5	-4	-7
Maximum	2	2	-1	8

Table 9.51
Positive and negative discrepancy values for each checklist, task 2



training. Initially 100 per cent of participants recorded fewer risk factors than the expert panel (up to -3). After training this transformed into 25 per cent correct risk identification, but a much wider spread of errors. Under-reporting dropped to only 13 per cent but all of these missed 7 or more risk factors. Similarly 13 per cent over estimated the risks by 7 or more factors. This wide diversity of scores suggests that a degree of confusion is generated by training, leading to a lack of precision.

The data were then split by checklist and by job position in order to examine the spread of positive and negative discrepancy values. Table 9.52 presents the descriptive statistics for checklist A by job position and Table 9.53 the corresponding statistics for checklist B. Each is presented for trained and untrained individuals.

Table 9.52
Positive and negative
discrepancy values
for checklist A task
2, by job position

	Discrepancy values (positive and negative values)			
	Team leader, line leader, line manager		Line worker, line operative	
	With training	Without training	Before training	Without training
Number of participants	2	5	10	8
Mean	-1.5	-1.4	-1.1	-1.3
Median	-1.5	-2	-1	-2
Mode	-2	-2	-4	-2
Standard deviation	0.7	2.5	2.4	2.2
Minimum	-2	-5	-4	-4
Maximum	-1	1	2	2

Table 9.53
Positive and negative
discrepancy values
for checklist B task
2, by job position

	Discrepancy values (positive and negative values)			
	Team leader, line leader, line manager		Line worker, line operative	
	With training	Without training	Before training	Without training
Number of participants	3	4	5	4
Mean	-2.7	0.75	-2	0.5
Median	-3	1	-2	0.5
Mode	-4	1	-3	-7
Standard deviation	1.53	0.5	1	6.14
Minimum	-4	0	-3	-7
Maximum	-1	1	-1	8

Figure 9.11 shows the percentile breakdown for the positive and negative discrepancy values by job position for checklists A and B. This shows that, for checklist A, line leaders accounted for a significant proportion of under-representation of risk factors before training. All of these individuals recorded a discrepancy value of between -1 and -2. After training this was reduced to 50 per cent under-reporting but 30 per cent were worse with -3 to -4 errors. The remaining line leaders either scored equal to the experts (20 per cent) or slightly over (up to +2 errors).

Line workers saw a similar amount of over estimation before and after training, but the 20 per cent who matched the expert score before training was translated into 75 per cent under-reporting risk factors with training. For this checklist there appears to be a negative effect from training such that this group fails to identify health hazards.

For checklist B, line leaders saw a similar trend. Before training all individuals under reported risk factors with discrepancy values of more than -1 (33 per cent up to -2, 67 per cent -3 to -4). With training this was converted into a significant over-awareness with 75 per cent of respondents identifying 1 or 2 more check items than were present.

Line workers suffered the greatest negative effect from training for checklist B. Initially all line workers under-estimated the number of risk items, with 40 per cent scoring 3 or 4 less than the experts. After training this converted into a wide spread of responses clearly split into quartiles from

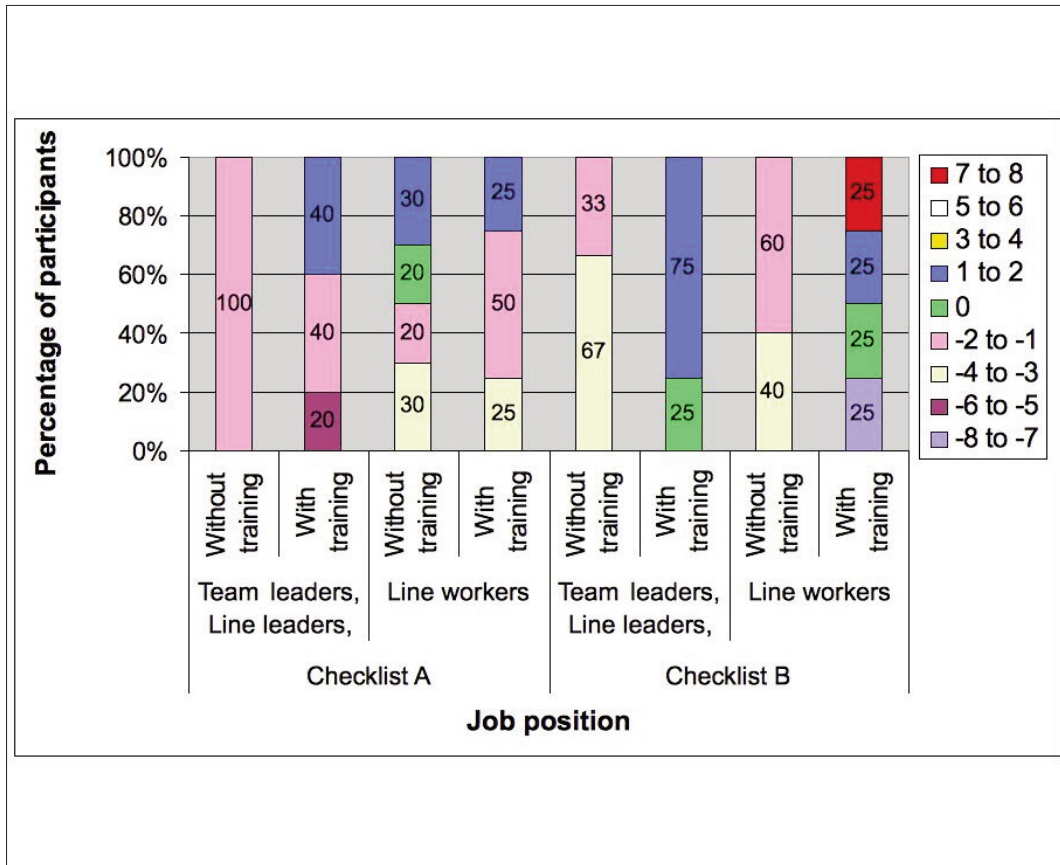


Figure 9.11 Percentage of participants' positive and negative discrepancy values for each checklist when used to assess task 2, by job position

–8 through to +8. This indicates that training for checklist B led to confusion in the line workers when appraising Task 2.

9.3.4 Overall risk level

The respondents' estimation of overall risk level was analysed to compare it with that recorded by the expert group. This is an important variable since the overall risk level category is likely to be the trigger for intervention, or at least to justify prioritisation of action. Accordingly, incorrect estimation of the overall risk could lead to failure to address potentially high risk activities or to waste resources addressing activities that are not particularly hazardous.

Task 1

When considering Task 1, Table 9.54 indicates the percentile agreement between participants and the expert scores for both checklists with and without training.

It can be seen that Checklist A generates a good correlation between participants and experts with 92 per cent correctly identifying the overall risk. Checklist B does not perform as well, with only half of the participants agreeing with the expert score. This increases to 70 per cent with training, which is positive, but still represents a significant error rate.

When comparison is made by job position and checklist variant the benefits of training for the different worker groups are apparent. Table 9.55 presents the percentile information for checklist A by job position. While line leaders lost some accuracy (20 per cent) after training, line workers gained almost as much (14 per cent). These results suggest that checklist A offers the possibility of achieving a good degree of precision in assessing overall risk, and that training is not necessarily required.

For checklist B (Table 9.56) the outcome is less clear. Line leaders maintained a reasonable precision before and after training, with 75 per cent matching the rating of the experts. However, line workers saw a significant increase in accuracy after training, rising from 33 to 67 per cent agreeing with the expert view. Most of the improvement was gained by elevating the assessment of those who had previously underestimated the overall risk. This would suggest that training with this checklist may have beneficial outcomes in ensuring higher risk activities are correctly identified.

Table 9.54
Percentage agreement of participants with the expert score for overall risk

	Percentage of participants			
	Checklist A		Checklist B	
Percentage of completed assessments that:	Before training	After training	Before training	After training
Agreed with the model response	92	92	50	70
Underrated the risk	0	8	20	10
Overrated the risk	0	0	10	20
No risk rating given	8	0	20	0

Table 9.55
Percentage agreement of participants with the expert score for overall risk using checklist A for task 1, by job position

	Percentage of participants			
	Team leader, line leader or line manager		Line worker or line operative	
Percentage of completed assessments that:	Before training	After training	Before training	After training
Agreed with the model response	100	80	86	100
Underrated the risk	0	20	0	0
Overrated the risk	0	0	0	0
No risk rating given	0	0	14	0

Table 9.56
Percentage agreement of participants with the expert score for overall risk using checklist B for task 1, by job position

	Percentage of participants			
	Team leader, line leader or line manager		Line worker or line operative	
Percentage of completed assessments that:	Before training	After training	Before training	After training
Agreed with the model response	75	75	33	67
Underrated the risk	0	0	33	17
Overrated the risk	0	25	17	17
No risk rating given	25	0	17	0

Task 2

A similar evaluation was undertaken for task 2. Table 9.57 shows the overall percentile agreement of the participants with the expert score for checklists A and B.

It can be seen that task 2 resulted in significantly worse correlation between participants and experts. For checklist A, only half of the participants correctly matched the overall risk score. This was raised somewhat by training, but only to 69 per cent, meaning that almost a third of participants erroneously assessed the task. Checklist B performed worse, with an initial correlation of only 38 per cent. This was also raised by training, but only to 50 per cent. In both instances the balance of responses was primarily an overestimation of the score, which could lead to wasted workplace resources.

A further evaluation splits these results by job position and presents them by checklist. Table 9.58 presents the percentile values for task 2 when assessed by checklist A, by job position.

	Percentage of participants			
	Checklist A		Checklist B	
Percentage of completed assessments that:	Before training	After training	Before training	After training
Agreed with the model response	50	69	38	50
Underrated the risk	0	0	0	10
Overrated the risk	50	31	50	20
No risk rating given	0	0	13	20

Table 9.57
Percentage agreement of participants with the expert score for overall risk

	Percentage of participants			
	Team leader, line leader or line manager		Line worker or line operative	
Percentage of completed assessments that:	Before training	After training	Before training	After training
Agreed with the model response	100	80	60	75
Underrated the risk	0	20	0	0
Overrated the risk	0	0	40	25
No risk rating given	0	0	0	0

Table 9.58
Percentage agreement of participants with the expert score for overall risk using checklist A for task 2, by job position

For this checklist, line leaders initially completely matched the expert score. However, with training, 20 per cent erroneously underestimated the risk level. Line workers, however, remained fairly consistent, with about two thirds of participants matching the expert score. Both before and after training the remainder overestimated the risk level.

For task 2 assessed by checklist B there is again greater diversity, as can be seen in Table 9.59. Line leaders seemed to respond to training well, raising their correct score from 33 to 80 per cent. Interestingly, the remainder before training overrated the risk level, whereas after training the balance underestimated it, suggesting a general downscaling of the risk perception.

	Percentage of participants			
	Team leader, line leader or line manager		Line worker or line operative	
Percentage of completed assessments that:	Before training	After training	Before training	After training
Agreed with the model response	33	80	40	20
Underrated the risk	0	20	0	20
Overrated the risk	33	0	60	40
No risk rating given	33	0	0	20

Table 9.59
Percentage agreement of participants with the expert score for overall risk using checklist B for task 2, by job position

Line workers fared worse, with training reducing an already low accuracy of 40 per cent to only 20 per cent. The majority of incorrect scores were overestimations of the risk presented. This suggests that checklist B may lead to line workers identifying a higher level of threat in some tasks, which while better than failing to correctly identify high risk activities, may waste safety resources. This situation seems unlikely to be addressed by training for this group.

9.3.5 Suggestions for improvement

A further analysis was undertaken of the responses relating to suggestions for improvement. This involved a scrutiny of both the number of participants who made suggestions as well as the number of suggestions they made. This helps to illustrate how effective the checklists may be in prompting the generation of interventions and the diversity of those interventions.

Task 1

Initially the evaluation only looks at task 1. Table 9.60 examines the number of suggested changes made by participants for the two checklists before and after training and presents the appropriate descriptive statistics. Checklist B generates a significantly larger mean number of suggested changes (9.9/9.1 versus 1.67/1.08), which may be as a result of the prompts within the checklist itself.

Table 9.60
Descriptive statistics for the number of suggested changes for each checklist when applied to task 1

	Number of suggested changes to reduce risks			
	Checklist A		Checklist B	
	Before training	After training	Before training	After training
Number of completed checklists	12	12	10	10
Mean	1.67	1.08	9.9	9.1
Median	1.50	0	6	3.5
Mode	0	0	0	0
Standard deviation	1.775	1.505	10.35	10.67
Minimum	0	0	0	0
Maximum	4	4	25	29

These data were further interrogated to establish the performance of the two checklists when compared by job position. Table 9.61 shows the results for Checklist A. It can be seen that this checklist resulted in a low number of suggested changes for both worker groups both before and after training. However, the line workers generated more suggestions than the line leaders in both instances.

Checklist B results are shown in Table 9.62. It can be seen that there are significantly more suggested changes generated than for checklist A for both worker groups whether trained or not. It is interesting to note that the line leaders nearly doubled the mean number of suggested changes after training, while the line workers reduced the mean number of suggested changes by approximately 25 per cent. This suggests that training has a meaningful effect with regard to generating intervention ideas for checklist B.

Suggestions for changes to activities from checklist A

The information for task 1 can be further evaluated by examining the number of participants who made suggestions or wrote more in-depth descriptions of changes. Table 9.63 presents this information for checklist A. It is clear that the number of participants (of both worker types) contributing in this way remains largely static before and after training. Changes in the number of suggestions after training are likely therefore be attributable to changes within individuals rather than altering the motivation of previously reticent participants.

The suggestions made by the individuals are included next.

	Number of suggested changes to reduce risks			
	Team leader, line leader, line manager		Line worker, line operative	
	Before training	After training	Before training	After training
Number of completed checklists	5	7	5	7
Mean	0.6	0.2	2.4	1.7
Median	0	0	3	2
Mode	0	0	3	0
Standard deviation	1.3	0.5	1.7	1.7
Minimum	0	0	0	0
Maximum	3	1	4	4

Table 9.61
Descriptive statistics for the number of suggested changes for checklist A when applied to task 1, split by job position

	Number of suggested changes to reduce risks			
	Team leader, line leader, line manager		Line worker, line operative	
	Before training	After training	Before training	After training
Number of completed checklists	4	4	6	6
Mean	4.5	8.5	13.5	9.5
Median	4.5	8.5	14.5	3.5
Mode	0	0	25	0
Standard deviation	3.87	9.26	12.1	12.4
Minimum	0	0	0	0
Maximum	9	17	25	29

Table 9.62
Descriptive statistics for the number of suggested changes for checklist B when applied to task 1, split by job position

	All participants		Team leader, line leader or line manager		Line worker or operative	
	Before training	After training	Before training	After training	Before training	After training
No. participants	6	5	1	1	5	4

Table 9.63
Number of participants who made their own suggestions or wrote more in-depth descriptions of changes that could be made (checklist A)

Team leaders, line leaders or line managers

- Before training
 - A stand for the pot could be used to avoid a wide grip being used to hold it. The bottle could be put into the fume cabinet for filling so as to avoid excessive shoulder/arm movement. Arm posture could be improved by moving the fume cabinet up/down for more support.
- After training
 - Some sort of tool to hold the heavy wide grip in place. This would remove factor 13, thus bringing the score down to 11 (low risk).

Line workers

- Before training
 - Empty contents of bottle which is being gripped into an easily accessible open container, therefore no gripping would be needed. Rather than twisting the body move body and feet to avoid awkward postures/movement. Place scales on raised platform to prevent neck being bent looking downwards, again avoiding awkward postures.
 - Lower opening to cabinet to allow better arm/shoulder posture. Smaller container to allow smaller grip of a holder for the container.
 - Not to do the task for too long. Change position in which you stand. Not to grip object for the whole length of the time performing the task. Put powder in a container that doesn't need to be held.
 - Have a wider neck on the bottles. Remove the glass that the worker is resting on or place some padding on it.
 - Clamp to hold the powder pot. Higher work surface. Smaller pot.
- After training
 - Place contents of bottle into smaller containers, ie cups which are lighter in weight using less force. Instead of twisting the body, move the feet and body, encouraging more movement and less static postures. Providing protection for the arms on the dust cabinet as it is uncomfortable. Place the scale higher so that the neck and back aren't bent over.
 - Opening hatch could be lowered. Smaller receptacle for powder in the hand. Raise the work cabinet.
 - Change the posture used during the task. Change the bottle in which powder is held so a wide grip is not needed. Have it placed on surface so it doesn't need to be held.
 - Holder for the powder container. Higher workstation within the fume cabinet to stop leaning on the perfect [sic] and neck craning.

While these suggestions are not very numerous they are of reasonable quality, identifying meaningful and practical interventions which would address the key check items of most concern. It should be remembered that the activities undertaken in task 1 are relatively technical compared to purely manual tasks. This would make it more challenging to identify appropriate interventions for participants who had not experienced this type of work activity.

Suggestions for changes to activities from checklist B

A similar analysis for checklist B reveals a similar result. A very low number of participants chose to contribute suggestions for changes although the numbers before and after training remain largely constant, as seen in Table 9.64.

Table 9.64

Number of participants who made their own suggestions or wrote more in-depth descriptions of changes that could be made (checklist B)

	All participants		Team leader, line leader or line manager		Line worker or operative	
	Before training	After training	Before training	After training	Before training	After training
No. participants	3	3	2	1	1	2

The actual suggestions for changes made by participants are shown next.

Team leaders, line leaders and line managers

- After training
 - Arms resting on edge of fume cupboard liable to cut off circulation.

Line workers

- After training
 - Protective glasses. Hair net. Smaller face mask. Better, neater gloves.
 - Change jobs and tasks every so often. Introduce machines to weigh the powder/put the powder into bottles. Make tables higher.

Task 2

Task 2 was a more conventional manual activity involving sorting flower stems and placing them onto a conveyor belt. This offers a range of potential health risks which do not require technical

appreciation. Interventions are likely to be more conventional and more universal, transferring more readily from other manual activities. Table 9.65 presents the descriptive statistics for the number of suggested changes made by participants when scrutinising task 2.

	Number of suggested changes to reduce risks			
	Checklist A		Checklist B	
	Without training	With training	Without training	With training
Number of completed checklists	12	16	8	9
Mean	2.1	0.81	10.38	3.56
Median	2.	0	11	0
Mode	1 and 2	0	0	0
Standard deviation	1.7	1.3	6.82	5.92
Minimum	0	0	0	0
Maximum	6	3	20	15

Table 9.65
Descriptive statistics for the number of suggested changes for each checklist when applied to task 2

More detail is revealed when the data are split by job position. Table 9.66 shows the descriptive statistics for checklist A. Here it can be seen that the line workers and line leaders generated similar levels of suggestions, suggesting that checklist A was equally effective at generating intervention ideas with these two groups and that training had little effect.

	Number of suggested changes to reduce risks			
	Team leader, line leader, line manager		Line worker, line operative	
	Without training	With training	Without training	With training
Number of completed checklists	2	5	10	8
Mean	3	1.2	1.9	0.88
Median	3	0	1.5	0
Mode	3	0	1 and 2	0
Standard deviation	0	1.6	1.9	1.3
Minimum	3	0	0	0
Maximum	3	3	6	3

Table 9.66
Descriptive statistics for the number of suggested changes for checklist A when applied to task 2, split by job position

A wider range of responses are noted for checklist B, as shown in Table 9.67. However, as with task 1, there is a drop in the mean number of suggestions made for both worker groups post training. Line leaders reduced to approximately half, while line workers dropped to nearly one tenth of the pre-training level. This would suggest that training has a deleterious effect on the generation of suggestions.

However, it is possible that these response changes can be largely accounted for by boredom, fatigue, lack of interest or other variables not directly related to the checklist structure. This may have real world implications if interest in the assessment process wanes during repeated activities. In this case then exposure to the assessment activity might need to be limited in order to maintain vigilance.

Table 9.67
Descriptive statistics for the number of suggested changes for checklist B when applied to task 2, split by job position

	Number of suggested changes to reduce risks			
	Team leader, line leader, line manager		Line worker, line operative	
	Without training	With training	Without training	With training
Number of completed checklists	3	5	5	4
Mean	9	5.4	11.2	1.25
Median	10	0	12	0
Mode	3	0	0	0
Standard deviation	5.57	7.47	7.98	2.5
Minimum	3	0	0	0
Maximum	14	15	20	5

Suggestions for changes to activities from checklist A

The suggestions made by the participants following use of checklist A are presented to add depth to the understanding of this participant role. Table 9.68 shows the number of participants who made suggestions. It can be seen that while the line leader contribution remained constant before and after training, line worker involvement decreased by almost two thirds.

Table 9.68
Number of participants who made their own suggestions or wrote more in-depth descriptions of changes that could be made (checklist A)

	All participants		Team leader, line leader or line manager		Line worker or operative	
	Before training	After training	Before training	After training	Before training	After training
No. participants	10	5	2	2	8	3

The suggestions for change made by the participants are presented below, split by job position. It can be seen that suggested changes are similar for both worker groups, which implies that the interventions are relatively obvious and more conventional than for task 1.

Team leaders, line leaders and line managers

- Before training
 - Change to another task more often. Make it easier to keep up. Change the layout of work area to stop the need to twist spine.
 - Bring holders and belt closer together. Raise belt and holder higher. More breaks.
- After training
 - Introduce microbreaks. Rearrange work space. Flowers are held in one hand continuously – perhaps a mechanical aid to hold flowers would also reduce the risk.
 - Main factor is holding flowers in one hand throughout the process. Perhaps shorten the length of time doing the job, and should monitor the situation. Environmental factors such as cold and wet could be improved.

Line workers

- Before training
 - The person working can also work in a high chair with back support, and reposition themselves at a tangent to the conveyor as this would prevent twisting the body.
 - More breaks.
 - Split the task and let take five minutes' rest every half hour. Suggest stretching and relaxing.
 - Slow down the line.
 - Less workload. Warmer clothes.

- Reduce cold/draughts, add isolated flooring, better protective clothing. Reduce overall duration of work between breaks. Reduce expected output of work. Work area could be improved in design to make it more ergonomic.
- Reduce amount of time individuals work at a task.
- Slower work rate would reduce neck movement. Better workstation layout would reduce movement. Worker could move right wrist more as it does not need to be static. A quick rest or job swapping could reduce time spent on continuous task. Better PPE could be supplied. Add more workers to reduce work rate.
- After training
 - Hold smaller bunches to encourage a power grip. When placing flowers on the belt move the feet and body rather than a twisting motion. Adjustable workstations so they are higher up.
 - Change posture. Have more breaks.
 - Resituate the conveyor so less twisting is required. Have a flower dispenser to eliminate the constant left hand grip.

Suggestions for changes to activities from Checklist B

Checklist B can be evaluated in the same way. Table 9.69 shows the number of participants who made suggestions, which are reproduced below.

	All participants		Team leader, line leader or line manager		Line worker or operative	
	Before training	After training	Before training	After training	Before training	After training
No. participants	4	1	1	0	3	1

Table 9.69

Number of participants who made their own suggestions or wrote more in-depth descriptions of changes that could be made (checklist B)

Team leaders, line leaders and line managers

- Before training
- After training

Line workers

- Before training
- After training
 - Better fitting gloves. Change area for better position to work in. Warmer clothing. Cold temperature. Draughty area. Headgear.

9.3.6 Training evaluation questionnaire

A questionnaire survey was undertaken on the completion of the training session to establish the attitudes of the participants. A brief summary of their responses are presented below.

Identification of risk factors

Participants from each of the companies were asked to identify up to six risk factors. Table 9.70 shows the mean value for each company. While the results appear optimistic, many of the risk factors given were not those recognised by conventional guides to MSD prevention. Accordingly this output should be considered encouraging inasmuch as risk factors were suggested but should be treated with caution in that the educational process had not been particularly effective in this regard.

Company	Mean number of risk factors identified
1	3.8
2	5.8
3	5.3
4	5.5

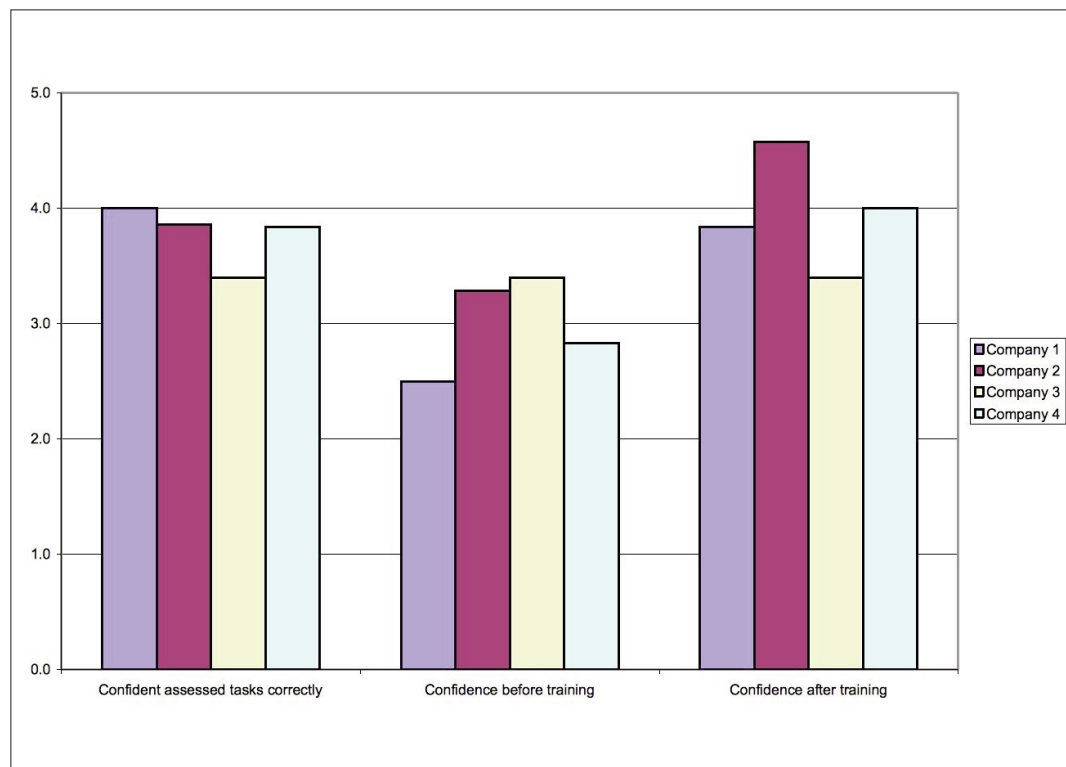
Table 9.70

Mean numbers of risk factors identified by participants at each company

Levels of confidence

Three questions were asked regarding the confidence of the individual. Firstly, the participants were asked how confident they were that had assessed the two tasks correctly (where 1 = not at all and 5 = very). They were then asked how confident they felt using the assessment tool before and after training. The mean results of this survey are shown in Figure 9.12. In this figure it can be seen that general confidence levels are high, with a significant increase of confidence after training for all except one group of company participants.

Figure 9.12
Mean confidence levels in assessment and use of tools



Training preferences

The participants were also asked about their training preferences. In particular they were asked to rate the following on a scale where 1 = strongly agree and 5 = strongly disagree:

- I would have liked more background information about the risk factors.
- I would have liked more information about how to complete the risk assessments.
- I would have liked more information about possible control actions or changes to make to reduce the risks.
- I think face-to-face training in the use of the assessments was more useful than following just the written instructions.
- I think following written instructions was more useful than attending a face-to-face training session.
- It was good to go through an example assessment with the trainer.

The mean results of the preferences are shown in Figure 9.13. It can be seen that the scores were similar across all four companies. Most notable was the preference for face-to-face training over the provision of written instructions, which indicates that training courses are likely to be well received. A worked example was also seen as a worthwhile approach. Overall, most participants seemed that they would have preferred more information on the background, the assessment tools and control actions.

Direct benefit

Participants were asked to provide feedback on three questions regarding positive benefit to themselves. These were:

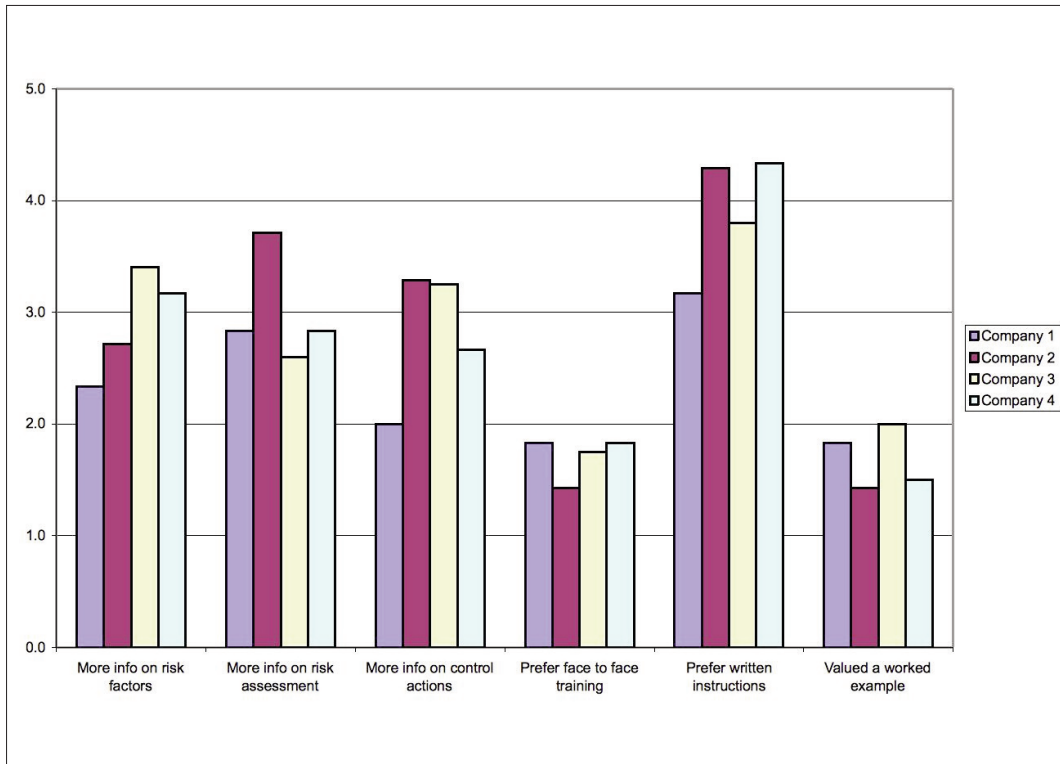


Figure 9.13
Participants' preferences following the training course

- Have you enjoyed the course?
- Has this course increased your knowledge about the risks for musculoskeletal problems?
- Has this course been of practical benefit to you? Will you use the knowledge you have gained back in the workplace?

These required simple binary responses (yes/no) and all respondents answered 'yes' to all three questions.

Overall opinion

Lastly, participants were asked for their overall opinion of four aspects of the course:

- the content of the course
- the course material
- the pace of the course
- the overall standard of the course.

Respondents were asked to answer these questions on a 1–5 scale, where 5 = excellent and 1 = poor. The mean responses for each company are shown in Figure 9.14. In each case the mean responses hovered around the values of 3 ('good') and 4 ('very good'). This appears to indicate that the training provided was considered of a reasonably high standard and considered acceptable by all the participants.

Figure 9.14
Participants' overall
opinion of the
training course

