Broad Lane, Sheffield, S3 7HQ
Telephone: +44 (0)114 289 2000
Facsimile: +44 (0)114 289 2500

HEALTH & SAFETY LABORATORY

Reliability and usability evaluation of the Manual handling Assessment Charts (MAC) for use by non-regulatory professionals

HSL/2003/19

D Lee and JJ Ferreira

Human Factors Group, HSL

CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

INTRODUCTION

The Manual handling Assessment Charts (MAC) is a new tool designed to help health and safety inspectors assess common risk factors in lifting, carrying and team handling operations (Monnington *et al.*, 2002). It was released to HSE and Local Authority inspectors on 5 November 2002 (HSE and HSL, 2002).

The Ergonomics Section of HSL was commissioned in January 2003 by the Human Factors Unit of HSE's Better Working Environment Division to evaluate the reliability and usability of the MAC when used by 100 non-regulatory health and safety professionals (non-inspectors).

The study also elicited feedback from 13 expert ergonomists including some of the leading ergonomics researchers, consultants and lecturers in the UK.

This report describes the reliability and usability of the MAC when used by non-inspectors, and provides recommendations for improvement.

MAIN FINDINGS

82% of participants reported that the MAC improved or greatly improved their confidence when assessing manual handling risks.

71% of participants reported that the MAC improved or greatly improved their understanding of manual handling risks

When asked to anticipate their future use of the MAC, only 3% reported seldom or never.

The overwhelming majority of users were very positive about MAC as a whole. Users liked the:

- Simplicity and ease of use
- Speed of use
- Intuitive colour scheme
- Traffic light pattern
- Step-wise approach
- Pictorial explanations
- Ability to determine which specific risk factors to focus prevention efforts on

Flaws were found, though, in both the reliability and usability of the MAC, and the study demonstrated that substantial improvements were both possible and necessary. Combining the quantitative and qualitative data provided clear direction in most cases for specific changes to the structural, descriptive and presentational aspects of the MAC

MAIN RECOMMENDATIONS

Thirty detailed recommendations are made in this report which can be summarised under the following four broad categories:

- Clearly define the limitations, legal standing and appropriate usage of the MAC.
- Expand the descriptive paragraphs to include additional colour coded tables, diagrams and explanatory text.
- Modify the score sheet design to minimise the need to flip between pages to record scores; to allow separate scoring of lifting, carrying and team handling operations; and to prevent users from scoring inappropriate risk factors.
- Adopt, test, and refine a single page format for the MAC flowcharts.

Changes based on seventeen of the recommendations were incorporated into the version of the MAC released to the public in August 2003 (HSE and HSL, 2003). Two of these had additional supporting documentation provided on HSE's website (http://www.hse.gov.uk/msd/) at the same time. Four were addressed solely through material made available on the website.

Acceptance of the remaining nine recommendations was deferred due to the need for further research and/or usability testing. The most significant deferred recommendations related to the current MAC format of separate charts for lifting, carrying and team handling and could only be properly addressed by combining the three flowcharts into one. Such an approach would directly simplify the risk assessment process, allow the lift, carry and team handling aspects of a single task to be assessed together, and alleviate the space constraints which are the root of many of the usability issues.

CONCLUSIONS

An iterative process has been employed since the initial development of the MAC to continually improve and refine its structure and presentation. The initial specification was for the development of a tool that would be suitable for use by health and safety inspectors who were not ergonomics specialists (Monnington *et al.*, 2002). Following release of this version (HSE & HSL, 2002) to inspectors in November 2002 a more recent focus has been on a desire to make it available for use by non-inspectors.

The results of this study indicate that although the version of the MAC released to inspectors demonstrates reasonable reliability and usability, significant improvements were necessary. Many of these changes have been incorporated in to the publicly available version of the MAC. Further research is necessary for the development of a single page MAC flowchart.

1 INTRODUCTION

1.1 BACKGROUND

The Manual handling Assessment Charts (MAC) is a new tool designed to help health and safety inspectors assess common risk factors in lifting, carrying and team handling operations (Monnington *et al.*, 2002). To ensure maximal suitability to the target users, the tool had to conform to the following criteria:

- It must be very quick and easy to use
- It must link in with other published information on manual handling, particularly that published by the Health and Safety Executive (HSE)
- It should intuitively indicate good manual handling practice.

A reliability study (Tapley, 2002) performed with a group of health and safety inspectors found that the MAC allowed inspectors to carry out a reliable initial identification, assessment and evaluation of significant manual handling risk factors. A usability study (Care *et al.*, 2002) used a two-phase questionnaire approach to obtain user feedback and recommend improvements for the version (HSE and HSL, 2002) released to HSE and Local Authority inspectors on 5 November 2002.

Some years prior to the development of the MAC, HSE had piloted a tool known as "Amber" (Dickinson *et al.*, 1999). This was designed to be used by inspectors as a hazard spotting tool when assessing manual handling activities by acting as a 'warning light' when a handling task was very hazardous to perform. The Amber chart suggested a load value that is considered to be acceptable to only one in ten of the workforce, after factors such as gender, action involved, distance for the hands away from the body and so on, are accounted for. This implies that actual loads above the Amber suggested value present a risk to 90% of the workforce.

The Amber chart was piloted over a three month period by seven HSE inspectors, who were first trained in its use. The feedback from the users suggested that there were too many problems associated with its use for it to be genuinely useful over and above the existing HSE guidance. In fact it was suggested that the relatively 'high' load values suggested by Amber might discourage some inspectors from pursuing action against employers in cases where the loads were smaller. Also, the constraints and limitations of the chart and the assumptions it required meant that it was not useful in too many cases. The development of the Amber chart, as an inspection tool was therefore discontinued.

The MAC, unlike Amber, achieved high levels of acceptability among HSE and LA inspectors. It was also apparent that the public, in particular non-regulatory health and safety professionals (NRPs) who had seen the MAC were keen to use it to assist them when carrying out manual handling risk assessments. There are inherent limitations in the use of the tool by NRPs as was never designed to meet the requirements for a full risk assessment but was intended as an aid to the inspectorate to help them identify high risk tasks which merited further investigation. Because the level of detail is greater than that provided by the risk filter in Appendix 1 of the HSE guidance on the MHOR (HSE, 1998), it falls between the filter and a full risk assessment.

A non-HSE study by Jones *et al.* (1999) had compared the ability of non-ergonomists and expert ergonomists to make manual handling risk assessments and implement changes in the health care setting. Trends in their data indicated that non-ergonomists appeared able to identify

hazards, though not necessarily to prioritise implementation tasks. With this in mind, it was felt that releasing the MAC to the public could be a useful aid to NRPs responsible for assessing manual handling risks, not only in identifying high risk manual handling operations but, more importantly, in prioritising improvement measures.

Therefore, the purpose of this investigation was to evaluate the reliability and usability of the MAC when used by non-regulatory health and safety professionals, a user group for whom it was not originally designed. Feedback from this group was seen as essential prior to the MAC being made available to the public.

1.2 AIMS

The overall aims of the study were:

- To evaluate whether non-regulatory professionals (NRPs) can reliably assess physical manual handling risk factors using the MAC.
- To obtain qualitative and quantitative feedback from NRPs on the usability of the MAC.
- To recommend improvements to the MAC prior to its release to the public.

1.2.1 Specific aims of the reliability test

The specific aims of the reliability test were:

- To assess levels of agreement between NRPs when using the MAC to score manual handling risk factors.
- To assess whether there are any differences in the assessment and scoring of MAC risk factors by NRPs who have received a briefing prior to testing (briefed) and NRPs who did not receive a briefing (non-briefed).
- To assess whether there are any differences in the assessment and scoring of MAC risk factors by expert ergonomists and other NRPs (non-ergonomists).
- To use the results of the reliability test to support any changes to the MAC prior to its release to the public.

1.2.2 Specific aims of the usability test

The specific aims of the usability test were:

- To assess whether there were any differences in the usability of the MAC between NRPs who have received a briefing (briefed) and NRPs who did not receive a briefing (non-briefed).
- To evaluate whether using the MAC helped to improve the confidence of NRPs when assessing manual handling risks as well as their understanding of manual handling risk factors.
- To obtain qualitative feedback from NRPs regarding use of the MAC and recommend improvements to the MAC in preparation for launch to the public.

2 METHODOLOGY

2.1 STUDY DESIGN

The study followed the design prescribed by Tapley (2002) in her reliability test of the MAC for health and safety inspectors. In addition, a questionnaire, based on that used by Care *et al.* (2002), was designed to elicit quantitative and qualitative feedback on the usability of the MAC.

2.2 SAMPLE GROUPS

Figure 1 shows the breakdown of MAC users, in particular non-regulatory professionals (NRPs), into sample groups.

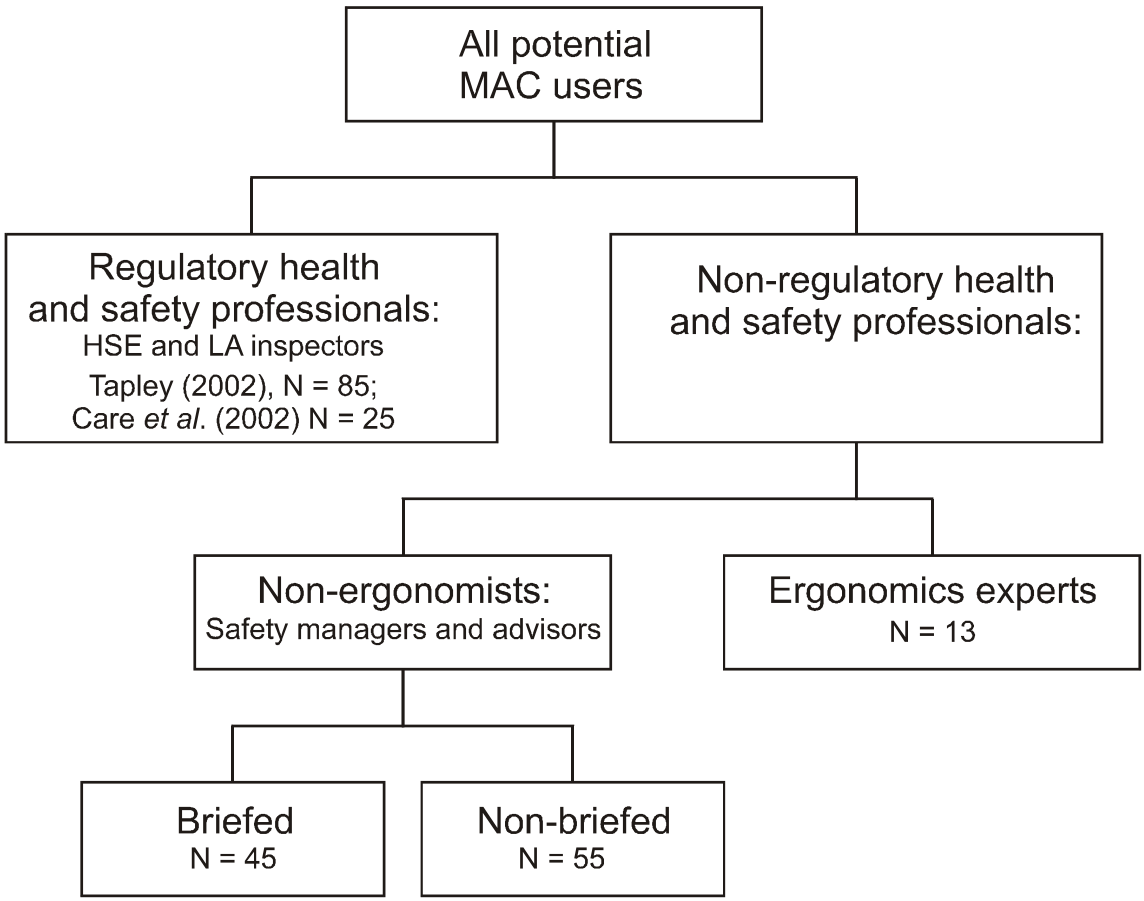


Figure 1 Breakdown of MAC users into sample groups

Non-experts were recruited with the help of industrial contacts, usually senior safety managers. These were asked to recruit staff from within their own organisations who had "some level of responsibility for assessing manual handling risks". However, the roles and experiences of users could not be controlled precisely and varied between organisations. The study encompassed users who worked in a variety of geographical locations in England, Scotland and Northern Ireland. The study also encompassed users in a variety of industries including: parcel shipment; finance; food and confectionary; foundry; furniture manufacturing; and tyre and rubber manufacturing.

In total, 13 MAC sessions were conducted, which served not only as a data collection exercise for the researchers, but also as a training opportunity for the users. There were 5 – 19 users at each session. Researchers assigned each session of non-expert testing to be either briefed (five sessions) or non-briefed (six sessions) in an attempt to create sample groups of similar size and to meet the expectations of particular industrial contacts. The other two sessions were expert sessions, conducted at two universities in the UK with ergonomics expertise.

Thus, the sample groups were:

- Briefed: Non-expert users who received a briefing about the MAC risk factors prior to participating in the reliability and usability tests (n = 45).
- Non-briefed: Non-expert users who did *not* receive a briefing about the MAC risk factors prior to participating in the reliability and usability tests (n = 55).
- Expert: Expert users who all received a briefing about the MAC risk factors prior to participating in the reliability and usability tests (n = 13).

2.3 SURVEY MATERIAL

To test reliability it was necessary to ensure that MAC users were assessing standardised tasks. Video footage from HSE inspections provided a convenient method of presenting manual handling operations and the associated risk factors to users in a controlled manner and it would not otherwise have been possible to expose all users to identical material. This method has the disadvantage that a risk assessor viewing a real task can view additional angles not provided by the video, and can solicit input from workers. The use of video may therefore have resulted in MAC users making decisions from incomplete or ambiguous information, thereby tending to decrease reliability artificially.

The video footage was previously used by Tapley (2002) and consisted of four tasks. Table 1 lists the tasks and the risk factors that were provided to users when the information could not be gleaned from the video. Some risk factors ("carry distance", "communication and co-ordination", and occasionally "floor surface") provided by Tapley (2002) were not provided in this study. This was because of need to investigate any usability problems that could emerge with these factors, despite the risk of reduced reliability from limitations in a video assessment.

Table 1 Description of survey material for each task and risk factors provided by researchers

Task	*Operation*	*Description*	*Risk factors provided*
1	Lifting	Operator handled 20 kg creels of wire onto spindles as part of a tyre making process.	Load weight / frequency Other environmental factors
2	Lifting	Operator lowered a 50 kg compost bale from a pallet to the edge of a hopper where it was split and fed into a flowerpot filling line.	Load weight / frequency Floor surface Other environmental factors
3	Carrying	Operator carried 20 kg tubs of pastry up steps to load into cookers.	Load weight / frequency Floor surface Other environmental factors
4	Team handling	Four operators lifted a 110 kg mould lid, turned it over and set it down on a trolley in the manufacturing of vehicle trim and matting.	Load weight Other environmental factors

2.4 STUDY PROTOCOL

The flowchart in Figure 2 shows the protocol followed at each session. In contrast to the reliability test of inspectors (Tapley, 2002), a discussion of scoring was performed only after all four tasks were assessed. The introductory briefing outlined the rationale for developing the MAC, its relation to the Health and Safety Commission's (HSC) priority programme for the prevention of work-related musculoskeletal disorders and the purpose of the study. The risk factor briefing included:

- An explanation of how to use the MAC.
- Video footage of task simulations that was used to outline the risk classifications for the more complex risk factors to assess: "load weight / frequency"; "hand distance from the lower back"; "vertical lift region"; "trunk twisting / sideways bending"; "asymmetrical trunk / load"; and "grip on load".
- Video footage of four additional tasks that was used to explain some of the risk factors in more detail. The video was paused to emphasise points if necessary and group discussion and interaction was encouraged.

Following the reliability test, each user completed a usability questionnaire (Appendix B). The questionnaire focussed on extracting concerns, confusions and recommendations from the users.

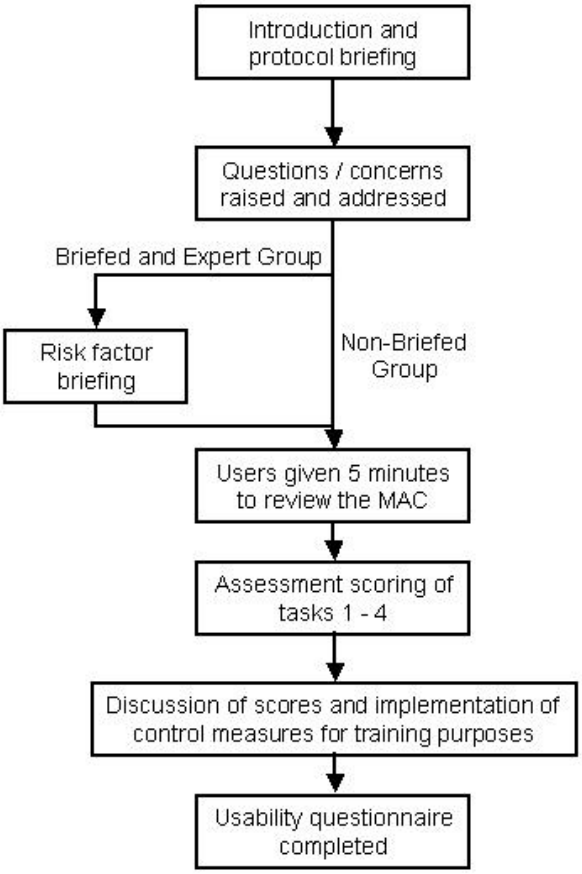


Figure 2 Flowchart of study protocol

2.5 DATA ANALYSIS

Statistical tests were performed using SPSS software.

2.5.1 Analysis of user data

User characteristics were analysed with descriptive statistics. The Mann – Whitney U test was used to compare users' experience in health and safety between the briefed and non-briefed groups.

2.5.2 Analysis of reliability data

As the MAC generates ordinal data, non-parametric statistical tests were used. Table 2 summarises the type of statistical analysis performed for each research question along with the null hypothesis.

Table 2 Statistical tests and null hypotheses for the reliability testing research questions

Research question	*Statistical test*	*Null hypothesis*
Do users within each sample group agree with each other?	Kendall's coefficient of concordance (W)	There is no significant agreement in the level of scoring.
How strong is the level of agreement for users within each sample group?	Kendall's coefficient of concordance (W)	There is no significant agreement in the level of scoring.
Is there a difference in risk factor scoring between the briefed and non-briefed groups?	Mann – Whitney U test	There is no significant difference in risk factor scoring between groups.
Is there a difference in risk factor scoring between the expert and non-expert users?	Mann – Whitney U test	There is no significant difference in risk factor scoring between groups.
Is there a difference in total scoring between any of the groups?	Kruskal – Wallis test	There is no significant difference between groups in the total scores.

Levels of scoring agreement were analysed with Kendall's test of concordance. The variables incorporated into this analysis included all of the risk factors relevant to the type of manual handling operation, as well as the final total score. Thus, for the lifting tasks (Tasks 1 & 2), levels of agreement were analysed for nine variables (eight risk factors + total score). For the carrying and team handling tasks (Tasks 3 & 4), levels of agreement were analysed for ten variables (nine risk factors + total score).

2.5.3 Analysis of usability data

Rating questions on the usability questionnaire generated ordinal data. Table 3 shows the statistical tests performed to answer each research question, as well as the null hypotheses. The qualitative data from the usability questionnaire were grouped into themes and analysed.

Table 3 Statistical tests and null hypotheses for the usability testing research questions

Research question	*Statistical test*	*Null hypothesis*
Is there a difference between the briefed and non-briefed group in the extent to which MAC usage improves user confidence when assessing manual handling risks?	Mann – Whitney U test	There is no difference in user confidence between groups.
Is there a difference between the briefed and non-briefed group in the extent to which MAC usage improves understanding of manual handling risks?	Mann – Whitney U test	There is no difference in understanding between groups.
Is there a difference between the briefed and non-briefed group in the extent to which the MAC helps determine what risk control action to take?	Mann – Whitney U test	There is no difference between groups in the extent the MAC helps determine risk control action.
Is there a difference in risk factor usability between the briefed and non-briefed groups?	Mann – Whitney U test	There is no difference in the usability of risk factors between groups.
Is there a difference in risk factor usability between the expert and non-expert groups?	Mann – Whitney U test	There is no difference in the usability of risk factors between groups.

3 RESULTS

3.1 USER DATA

3.1.1 User roles

Of the 100 non-ergonomists who participated in the study, 45% were briefed and 55% were non-briefed. Table 4 summarises the roles they performed for their employers as NRPs.

Table 4 Numbers (percentages) of users in various health and safety roles

Health and safety role	*Briefed*	*Non-briefed*	*Total*
Health and Safety Manager	17 (37.8%)	13 (23.6%)	30 (30%)
Health and Safety Trainer or Advisor	7 (15.6%)	12 (21.9%)	19 (19%)
Health and Safety Rep. or Team Member	9 (20%)	8 (14.6%)	17 (17%)
Line Manager with health and safety responsibilities	6 (13.3%)	6 (10.9%)	12 (12%)
Occupational Health Staff	3 (6.7%)	7 (12.7%)	10 (10%)
Industrial Engineer	0	7 (12.7%)	7 (7%)
Unspecified	3 (6.6%)	2 (3.6%)	5 (5%)
Total	45	55	100

3.1.2 User experience

Figure 3 categorises the experience of users within the briefed and non-briefed groups.

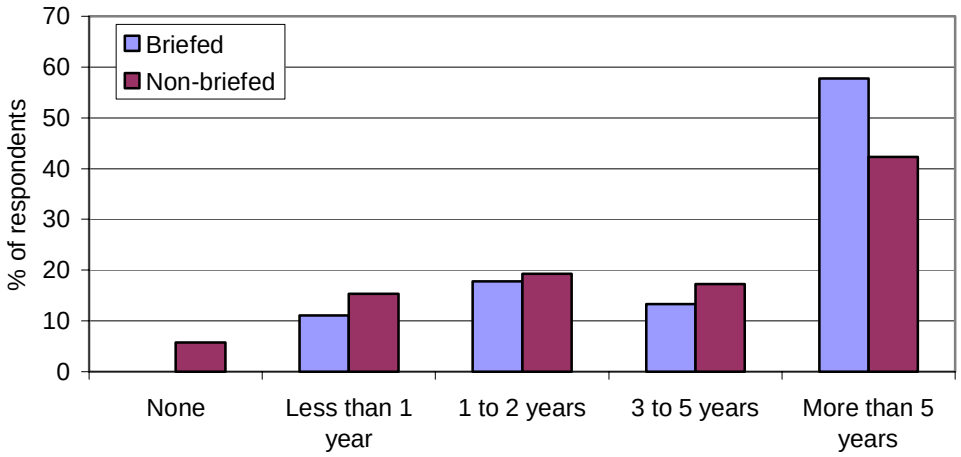


Figure 3 User experience in health and safety

Approximately 50% of all users had more than five years of experience in health and safety. Although the briefed group appears slightly more experienced than the non-briefed group, the difference was not found to be significant ($p > 0.05$). 76% of briefed users and 75% of non-briefed users reported previous experience in performing manual handling risk assessments.

3.1.3 Expert user roles

Expert feedback was elicited from 13 ergonomists. Twelve experts performed a range of ergonomics activities including research, lecturing and consulting. There was one ergonomics student within the group.

3.1.4 Expert user experience

Nine expert ergonomists had more than five years experience assessing manual handling risks; one expert had 3 – 5 years experience and two experts had 1 – 2 years experience. One expert did not have any experience in the area of manual handling.

3.2 RELIABILITY RESULTS

3.2.1 Levels of scoring agreement

Table 5 shows the levels of scoring agreement amongst briefed, non-briefed and expert groups for each task, as reported by Kendall's coefficient of concordance (W). This can range from 0 (no association) to 1 (perfect association). Although no exact values constitute strong, moderate or poor associations, levels of agreement measured with W can be interpreted in the following manner (MG Boocock, personal communication):

- Poor association: $0.0 < W < 0.6$
- Moderate association: $0.6 < W < 0.7$
- Good association: $0.7 < W < 0.8$
- Strong association $0.9 < W < 1.0$

On the basis of these interpretations, Table 5 indicates a good association in scoring for the lifting tasks (Task 1 and 2). There was a moderate association in scoring for the carrying task (Task 3) and a poor-to-moderate association in scoring the team handling task (Task 4). For all tasks, there were significant ($p < 0.05$) associations in scoring within the briefed, non-briefed and expert groups.

Table 5 Levels of agreement within briefed, non-briefed and expert groups measured with Kendall's coefficient of concordance (W)

Task	*Briefed (n = 44)*	*Non-briefed (n = 55)*	*Expert (n = 12)*
1	0.701	0.820	0.789
2	0.822	0.820	0.789
3	0.660	0.705	0.734
4	0.590	0.656	0.602

Although this test does not compare the three groups, levels of agreement between the groups seem similar. The non-briefed group had slightly higher levels of agreement than the briefed and expert group. Comparisons between groups were made for individual factors using the Mann-Whitney test and the Kruskal-Wallis test for total scores (Table 2)

3.2.2 Total score results

For each task summary statistics are listed of the total scores of users in the briefed, non-briefed and expert groups (Tables 6 – 9). Group sizes vary for each task because total scores were only calculated if users assessed all of the required risk factors. Kruskal-Wallis tests show that there were significant differences ($p < 0.05$) in the total scores between groups on both Task 1 and Task 4. Inspection of the group means for Task 1 showed that the non-briefed group scores were lower than the non-briefed and expert group scores. The non-briefed group had significantly lower scores for the "hand distance from the lower back", "postural constraints", "grip on load" and "other environmental factors" risk factors, which would explain the difference in the total scores. For Task 4 the briefed group scores were higher than for both the non-briefed and expert groups, with significantly higher scores for the "vertical lift region", "trunk twisting / sideways bending" and "floor surface" risk factors.

Table 6 Task 1 scoring summary

Group	*N*	*Mean*	*Std deviation*	*Range*
Briefed	44	14.28	3.43	7 – 22
Non-briefed	55	12.14	2.53	7 – 19
Expert	12	14.22	2.91	10 –18

Table 7 Task 2 scoring summary

Group	*N*	*Mean*	*Std deviation*	*Range*
Briefed	45	22.46	2.91	14 – 29
Non-briefed	55	22.69	2.56	17 – 27
Expert	10	20.00	6.84	3 – 26

Table 8 Task 3 scoring summary

Group	*N*	*Mean*	*Std deviation*	*Range*
Briefed	42	11.90	2.95	7 – 25
Non-briefed	48	12.17	2.70	6 – 18
Expert	11	12.67	3.24	9 – 17

Table 9 Task 4 scoring summary

Group	*N*	*Mean*	*Std deviation*	*Range*
Briefed	43	14.26	4.39	7 – 25
Non-briefed	47	11.64	3.22	6 – 18
Expert	9	11.78	2.28	9 – 17

3.2.3 Scoring differences between briefed and non-briefed groups

Significant differences ($p < 0.05$) between the briefed and non-briefed groups were reported for eight of the twelve risk factors. These included: "load weight / frequency", "hand distance from the lower back", "vertical lift region", "trunk twisting / sideways bending", "postural constraints", "grip on load", "floor surface" and "other environmental factors". The significant differences found highlight some usability problems with specific risk factors and general features of the MAC. They may also reflect limitations of the video-based methodology as well as confounding sample differences. There were no significant differences found in the risk factors listed solely for carrying or team handling. For each risk factor, z-scores and p-values are reported in Appendix A.

3.2.4 Scoring differences between expert and non-expert groups

A significant difference ($p < 0.05$) between expert and non-expert groups was found in the scoring of the "hand distance from the lower back" and "grip on load" risk factors. For the "vertical lift region", differences between experts and non-experts also approached significance ($p = 0.05$). For each risk factor, z-scores and p-values are reported in Appendix A.

3.3 USABILITY RESULTS

3.3.1 Influence of the MAC on user confidence when assessing manual handling risks

Figure 4 shows the extent to which briefed and non-briefed users felt that MAC usage improved confidence when assessing manual handling risks. In total, 29% reported that the MAC "greatly improved" their confidence when assessing manual handling risks. 53% reported that the MAC "improved" their confidence, while 15% reported it "slightly improved" their confidence.

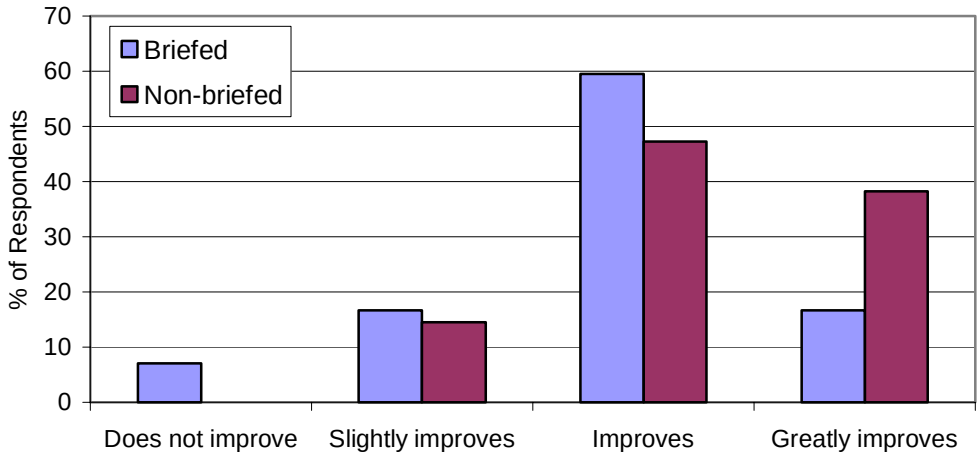


Figure 4 Perception of the extent to which MAC usage improves the confidence of users when assessing manual handling risks

The extent to which the MAC was perceived to improve user confidence was significantly greater ($p < 0.05$) for the non-briefed group than the briefed group (Figure 4). Often during the briefing, some users highlighted the recognised limitations of the MAC as a complete manual handling assessment. As such, the briefed users became aware that using the MAC alone would not be sufficient for their most complex manual handling tasks.

3.3.2 Influence of the MAC on user understanding of manual handling risk factors

Figure 5 shows the extent to which briefed and non-briefed users felt that MAC usage improved their understanding of manual handling risks. After using the MAC, 19% of users thought their understanding of manual handling risks "greatly improved", 52% thought their understanding "improved", and 20% felt that their understanding "slightly improved". The extent to which MAC usage improved the understanding of manual handling risk assessment would depend upon the prior knowledge and experience of users. However, the results suggest that that the MAC has great potential as a training tool for health and safety managers. This was often mentioned during the site visits. As there was no significant difference between briefed and non-briefed users, simple distribution of the MAC should improve the public's understanding of manual handling risk factors at the workplace and during daily activities.

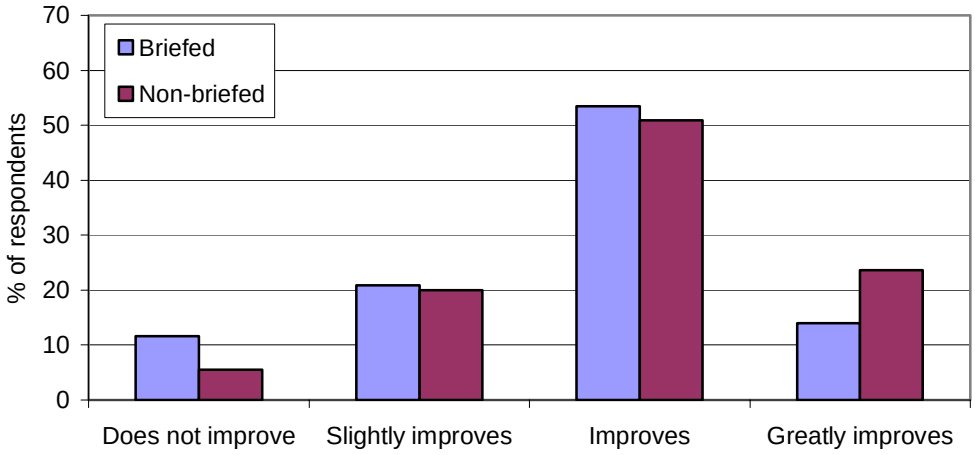


Figure 5 Perception of the extent to which MAC usage improves the understanding of users when assessing manual handling risks

3.3.3 MAC assistance in determining the action to take

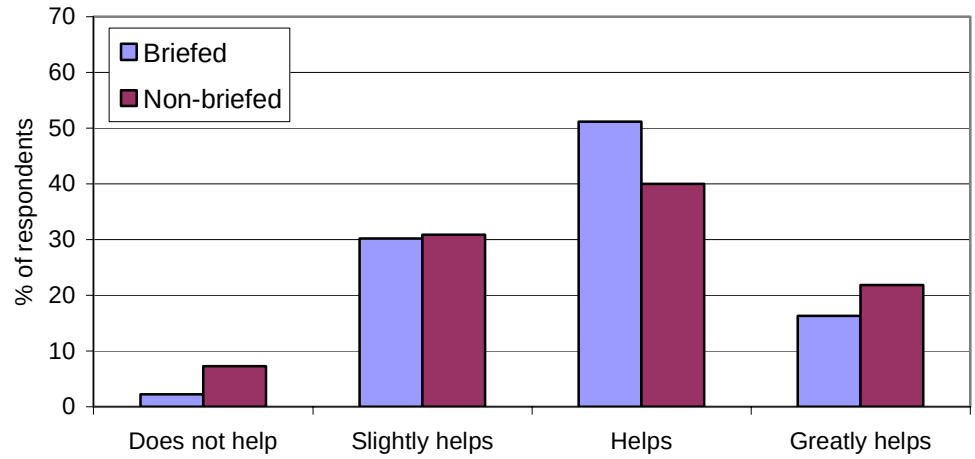


Figure 6 Perception of the extent to which MAC helps to determine the action to take to reduce the manual handling risk

Figure 6 shows the extent to which briefed and non-briefed users felt that the MAC helped determine the action to take to reduce manual handling risks. In total, 19% indicated that the MAC "greatly helps", 45% indicated that it "helps", and 31% indicated that it "slightly helps". There was no significant difference between briefed and non-briefed users.

3.3.4 Anticipation of future MAC usage

When asked to anticipate their future use of the MAC, 36% of users responded "almost always", 30% responded "often", 40% responded "occasionally", while only 3% responded "seldom or never" (Figure 7). Non-briefed users anticipated that they would use the MAC significantly more than briefed users ($p < 0.05$). Again, during the briefing discussion, users became aware of some manual handling situations, such as pushing and pulling, for which the MAC would not be appropriate. This may have influenced their response in comparison to the non-briefed group.

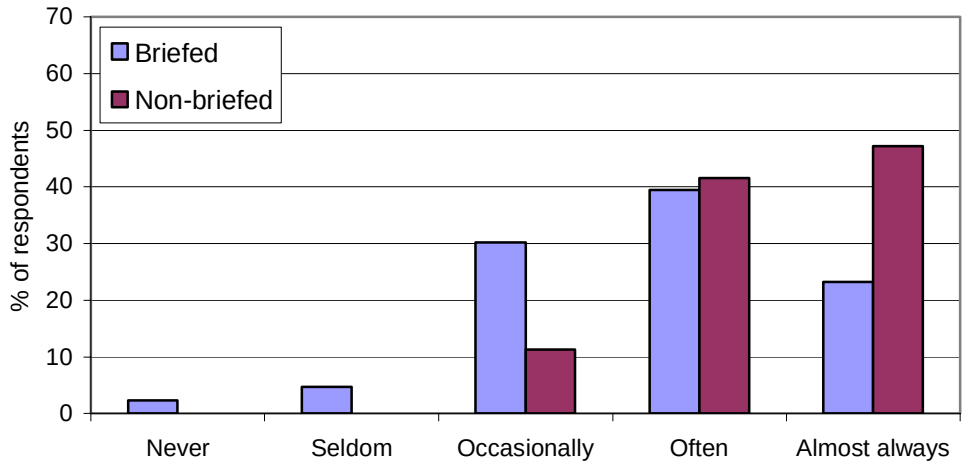


Figure 7 Anticipated frequency of future usage of the MAC

3.3.5 Usability rating summary of the MAC risk factors

The percentage of all users (briefed, non-briefed and expert) who rated the usability of each risk factor favourably is shown in Figure 8. A favourable response was defined as a score of 1 or 2 on a 5-point scale, with 1 representing "very easy" and 5 representing "very difficult". For all risk factors, more than 50% of users rated the usability favourably. The "load weight / frequency" and "obstacles en route" risk factors received the most favourable ratings. Less usable risk factors, receiving favourable ratings from 50 – 60% of users included: "hand distance from the lower back", "other environmental factors" and "communication and co-ordination".

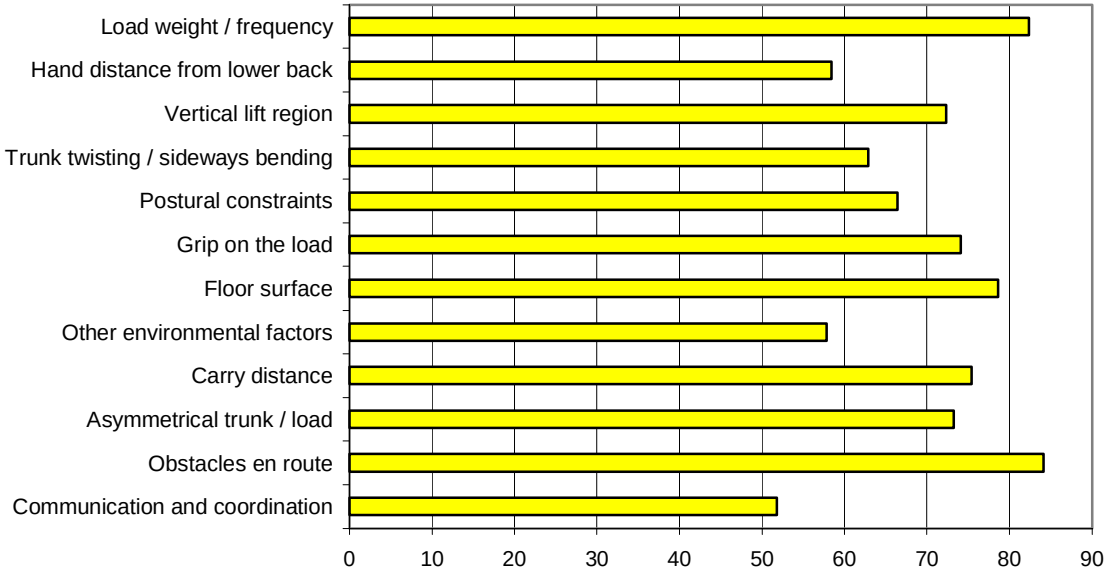


Figure 8 Percentage of all users who indicated a favourable usability rating for MAC risk factors

3.3.6 Differences in usability rating between briefed and non-briefed groups

Significant differences ($p < 0.05$) were found between briefed and non-briefed users in the usability rating of three risk factors. These included: "vertical lift region", "trunk twisting / sideways bending" and "other environmental factors". For each risk factor, z-scores and p-values are reported in Appendix A.

3.3.7 Differences in usability rating between expert and non-expert groups

Significant differences ($p < 0.05$) were found between expert and non-expert users in the usability rating of four risk factors. These included: "hand distance from the lower back", "carry distance", asymmetrical trunk / load" and "communication and co-ordination". For each risk factor, z-scores and p-values are reported in Appendix A.

4 DETAILED EXAMINATION OF RISK FACTORS

4.1 LOAD WEIGHT / FREQUENCY

4.1.1 Levels of scoring agreement

Levels of scoring agreement of load and frequency were good, both within and between groups, with more than 86% of users scoring the same level of risk for each task (Figure 9). For the 50 kg bale handling task (Task 2), a significantly greater proportion ($p < 0.05$) of briefed users rated the "load weight / frequency" risk factor red, rather than the correct level, purple. This may be due to uncertainty regarding the applicability of the MAC 'lifting' chart to 'lowering' tasks. On the video, it did not appear as though the operator was applying as much force to lower the bale as would be required to raise it to its original pallet height. The MAC does not state whether lowering tasks should also be assessed in the same manner as lifting tasks. This may also account for one expert scoring the risk factor as green.

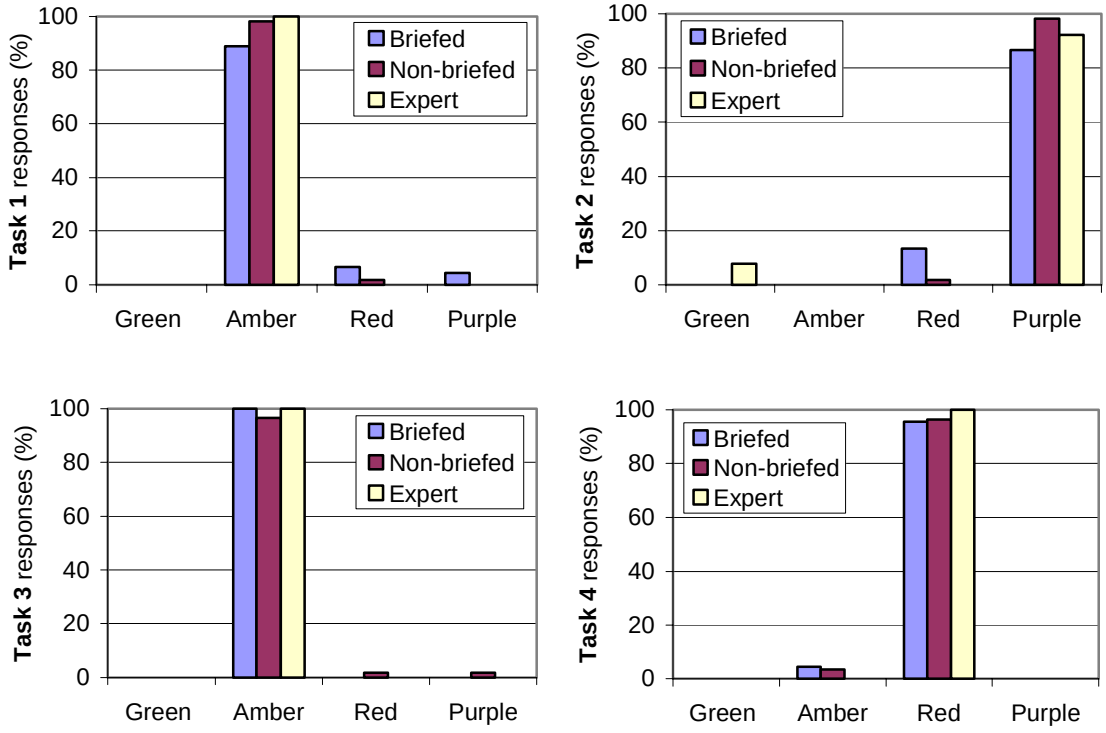


Figure 9 Inter- and intra-group agreement for the "load weight / frequency" risk factor scores

4.1.2 Usability

The "load weight / frequency" risk factor was judged to be the most usable of all risk factors. 84% of the briefed group, 84% of the non-briefed group and 69% of the expert group rated the usability of the "load weight / frequency" favourably, indicating a 1 or 2 on a 5-point scale (Figure 10). There was no significant difference in the usability rating between the briefed and non-briefed groups, which suggests that the load weight / frequency graphs are intuitive.

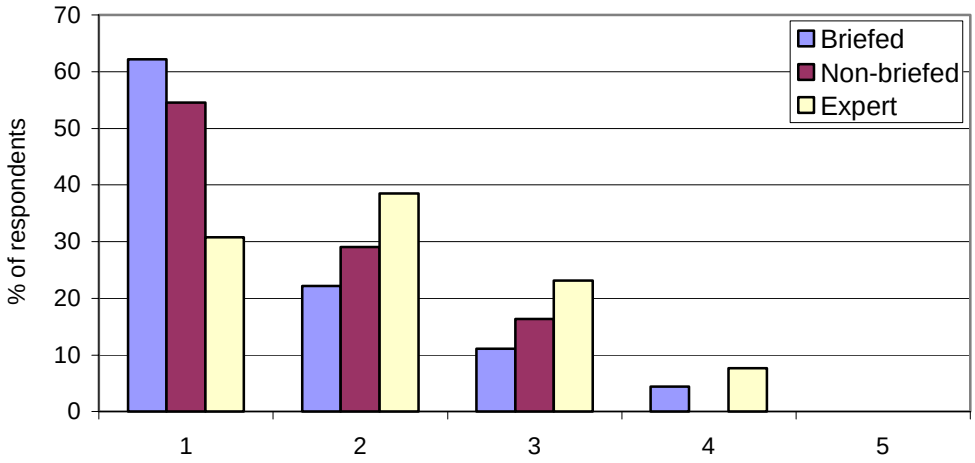


Figure 10 Usability ratings of the "load weight / frequency" risk factor (1 = Very easy, 5 = Very difficult)

Poorer usability ratings reflected some users' concerns that the MAC appears to require a constant load and frequency, with no guidance provided on how to assess tasks with variable weights and processes. Two users felt that the x-axis labelling was not intuitive, while another thought that frequency could be expressed better as a series of ranges.

4.2 HAND DISTANCE FROM THE LOWER BACK

4.2.1 Levels of scoring agreement

Levels of agreement for the "hand distance from the lower back" scores were quite variable (Figure 11). In three of the four tasks, the non-briefed group showed approximately 80% agreement. However, levels of agreement, at 41% to 80%, were unexpectedly lower for the briefed and expert groups. Usability data suggests that this may be due largely to the subjective interpretation that is required to assess the extent of shoulder (upper arm) and trunk flexion. Users were shown footage of a fully stacked pallet in Task 2, and were asked to project ahead and produce an assessment for the whole task rather than just unloading the top of the pallet, as seen on video. However, some users still only assessed the components of the task they saw on the tape. Also, the loss of stereoscopic vision and corresponding loss of visual detail when viewing videotape may have reduced users' ability to assess accurately (Troup and Rauhala, 1987). This may also be the case in Task 3, where the camera angle hindered postural assessment in the sagittal plane during the step climb.

For Task 1 (creel handling), significantly more briefed than non-briefed users assessed the "hand distance from the lower back" as red ($p < 0.05$). This may be due to the researchers' emphasis during the briefing that the 'worst case scenario' should be assessed. There was a significant ($p < 0.01$) difference in scoring between expert and non-expert groups for Task 2 that may reflect prior experiences of the experts who were used to assessing this risk factor according to alternative criteria. Some of the experts were previously involved with the development of other manual handling and postural assessment tools for ergonomists.

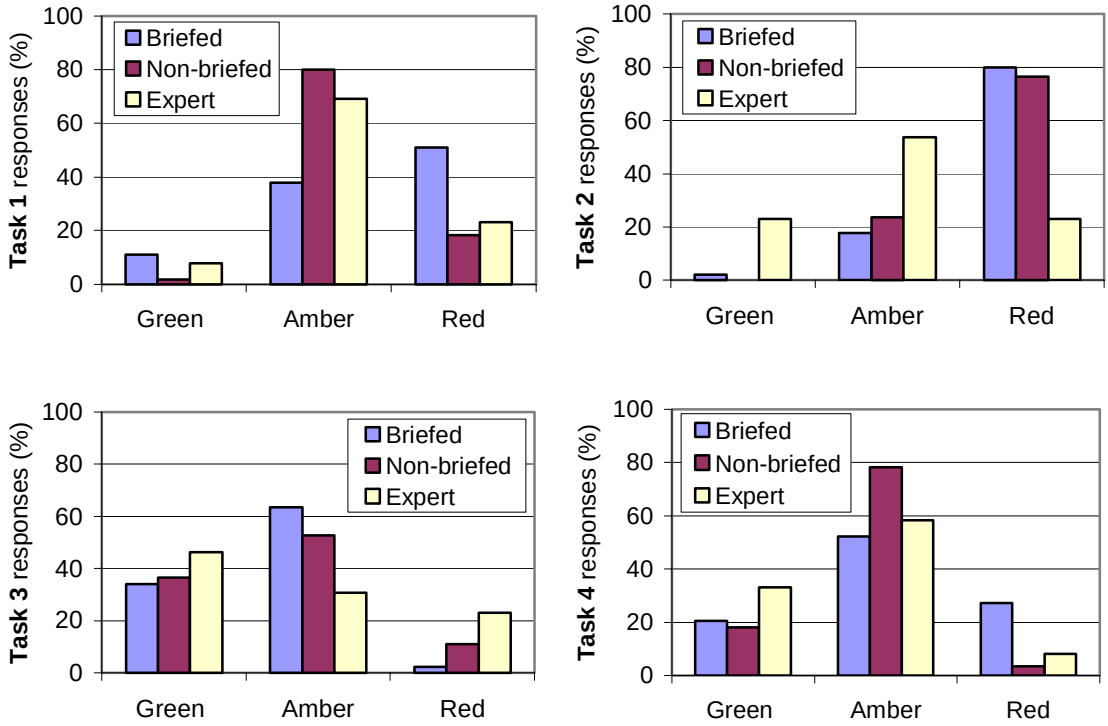


Figure 11 Inter- and intra-group agreement for the "hand distance from the lower back" risk factor scores

4.2.2 Usability

Figure 12 shows that 70% of the briefed group and 55% of the non-briefed group rated the usability of the "hand distance from the lower back" risk factor favourably, scoring 1 or 2 on a 5-point scale. The expert group rated the usability to be significantly more difficult ($p < 0.01$), possibly due to their familiarity with previously used criteria.

Difficulties were expressed in differentiating the boundaries of the risk categories. This is to be expected in any risk assessment tool that requires users who lack extensive training or practice to make a series of subjective judgements. However, some of the difficulty can also be explained by ambiguous, imprecise or conflicting information in the MAC.

Those with ergonomics expertise recognise that the "hand distance from the lower back" risk factor was intended to assess the horizontal reach component. Those users who took the wording literally felt that as the hands move farther from the lower back (including movement in the vertical plane) the risk score should increase. This results in double scoring of the vertical component. Furthermore, the coloured boxes suggest that risk is dependent on the angle of the upper arm with respect to the body, whereas the pictures suggest that risk is dependent on the horizontal distance of the hands from the low back. In some situations, these three interpretations will lead to the same result, while in other situations, three different risk ratings may result.

Some users suggested that usability could be enhanced with further guidance on how to assess variation in the "hand distance from the lower back" due to variable product type and product placement.

The MAC does not mention the angle at which the trunk or upper arms must be positioned before they are classified as 'angled'.

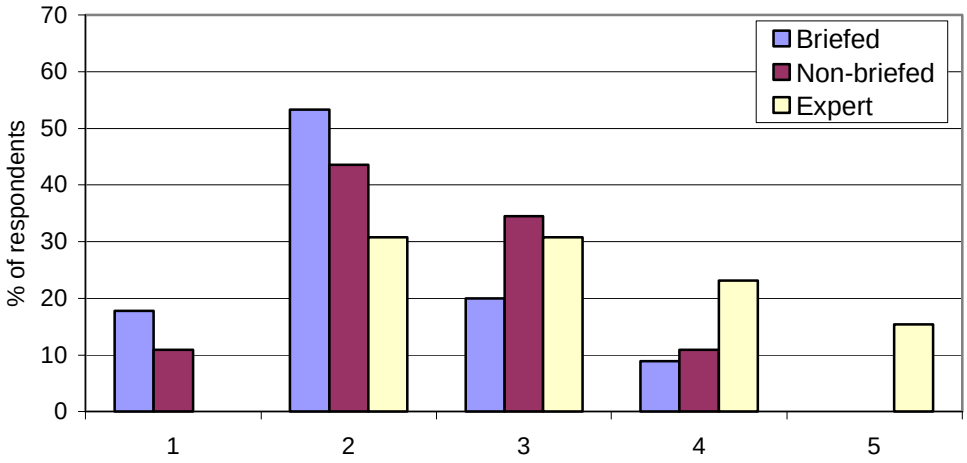
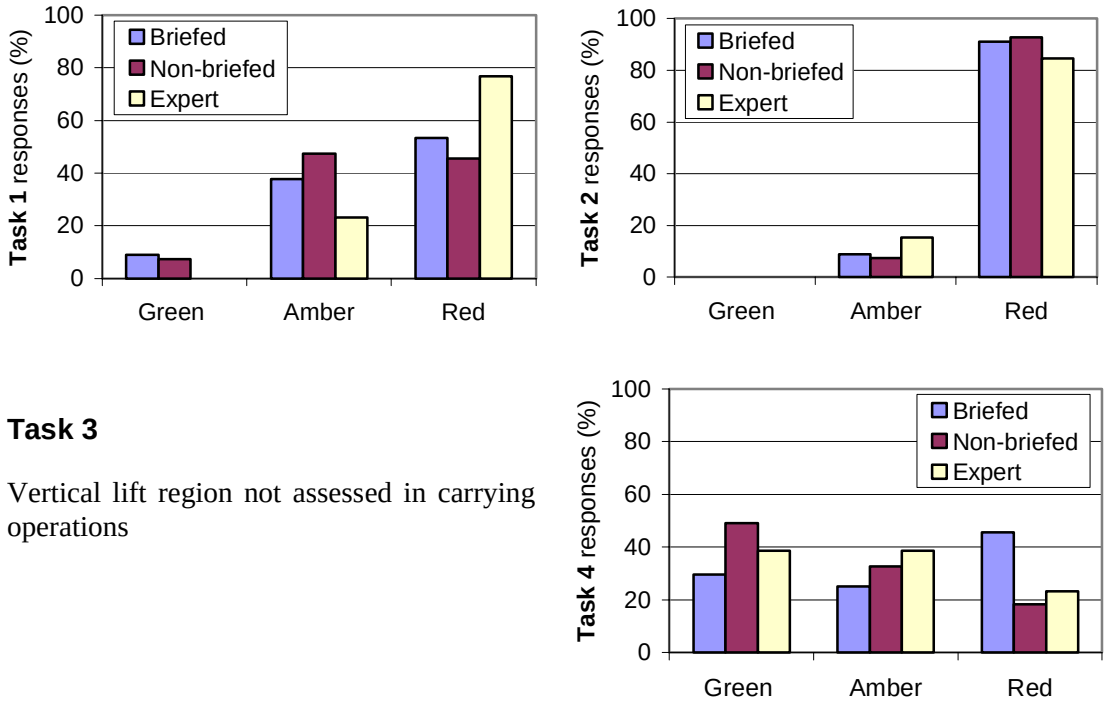


Figure 12 Usability ratings of the "hand distance from lower back" risk factor (1 = Very easy, 5 = Very difficult)

4.3 VERTICAL LIFT REGION

4.3.1 Level of scoring agreement

The "vertical lift region" showed a high level of agreement for Task 2, when the load was clearly lifted from above head height (Figure 13). However, lower levels of agreement (40% – 50%) for both Task 1 and Task 4 suggested a problem with the description of this risk factor.



Task 3

Vertical lift region not assessed in carrying operations

Figure 13 Inter- and intra-group agreement for the "vertical lift region" risk factor scores

For these tasks, some users assessed the vertical position from which the load was lifted initially. This procedure follows the description within the lifting and team handling flowchart boxes; for example, "lift from floor level or below…" However, the text instructs users to "observe the position of the load at the start of the lift and as the lift progresses". This discrepancy of instruction was thought to explain a significant difference ($p < 0.01$) between the briefed and non-briefed groups for Task 4, where the load is initially lifted at waist level (green), but then moved to approximately shoulder/head height (amber/red) as the lift progresses. Further discrepancies occurred for Task 4 because there was debate as to whether the load was supported by the trolley when the operators' hands moved to shoulder/head height.

4.3.2 Usability

Figure 14 shows that 87% of the briefed group, 63% of the non-briefed group and 62% of the expert group rated the usability of the "vertical lift region" risk factor favourably (1 or 2 on a 5-point scale). There was a significant difference ($p < 0.01$) in the usability rating between briefed and non-briefed groups that may reflect the additional instruction provided by researchers during the briefing to assess the "vertical lift region" as the lift progresses, and not just the initial lifting position.

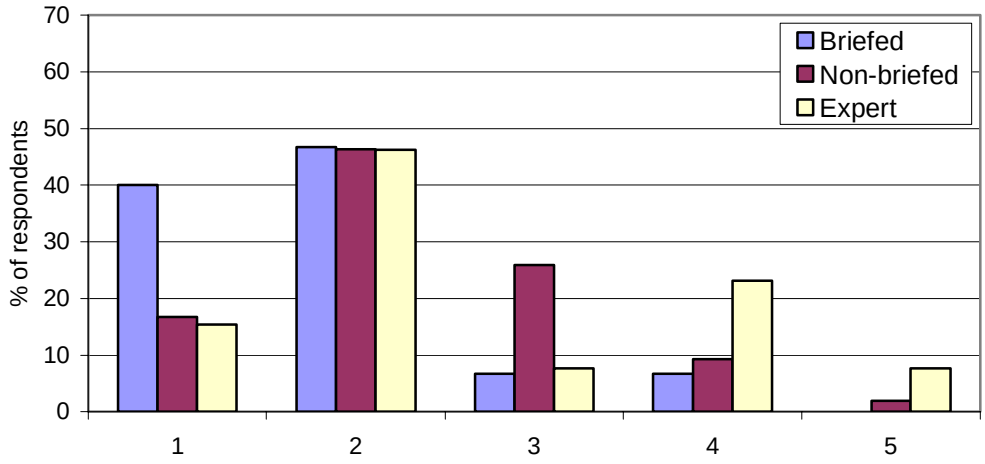


Figure 14 Usability ratings of the "vertical lift region" risk factor (1 = Very easy, 5 = Very difficult)

Qualitative feedback indicated that some users experienced difficulty differentiating the vertical lift region boundaries, particularly with respect to whether the position of the hands or the position of the load should be assessed. Some users also predicted potential confusion for a lift from floor level to above head height. As the red level of risk for the "vertical lift region" represents both positions, there was confusion as to whether this risk factor should be double-scored in this situation.

Some users found the "vertical lift region" description to be confusing or misleading. The reason cited was that the text makes reference to "the start of the lift and as the lift progresses," while the coloured boxes use the words "lift from," which implies consideration of just the starting point, not the progression through the lift. Guidance on how to assess highly variable activities was also requested.

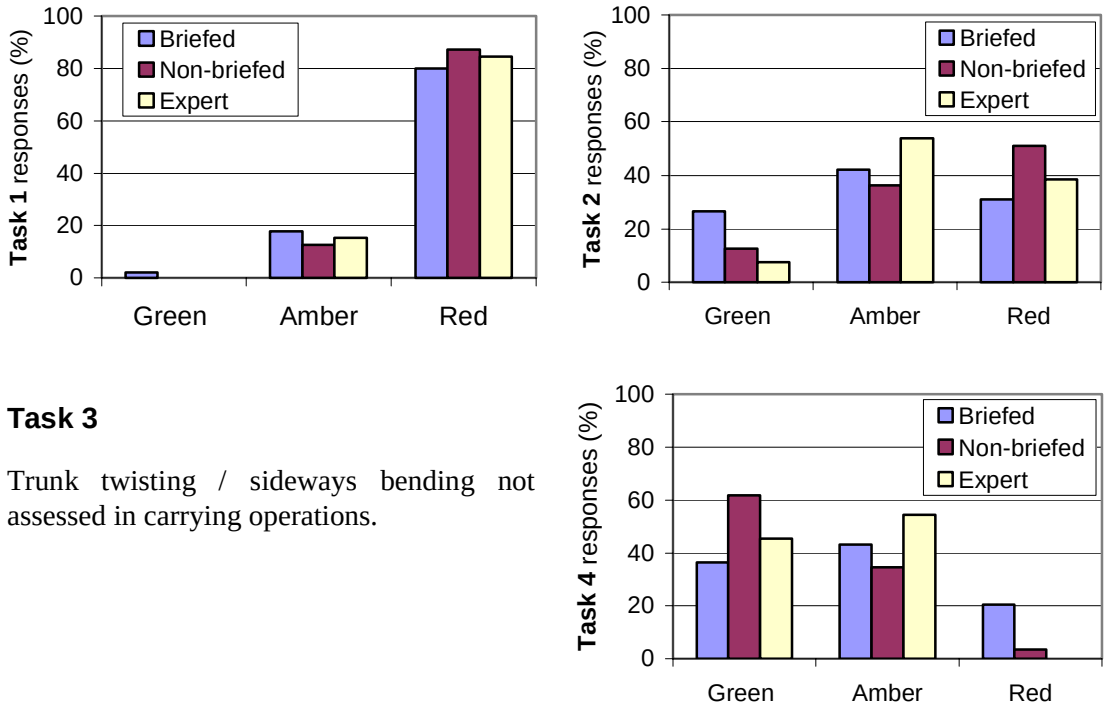
One expert questioned whether lifting from below knee-level possesses enough risk to warrant a red, high-risk classification. Alternatively, some industry representatives felt strongly that lifting from the floor should not be considered a high-risk activity as they felt this risk can be

eliminated by training in good techniques. The researchers are not aware of any scientifically valid data that supports this opinion.

4.4 TRUNK TWISTING / SIDEWAYS BENDING

4.4.1 Levels of scoring agreement

Figure 15 shows that unless trunk twisting and sideways bending were quite pronounced, as in Task 1, the variation in the scoring this risk factor was high. For Tasks 2 and 4 there was only 40 – 60% agreement for the briefed, non-briefed and expert groups. As with the "hand distance from the lower back" risk factor, much of the variation may be due to whether users recognised clues in the videotape to project ahead and produce an assessment for the whole task rather than just the video evidence. In addition, there were significant differences between the briefed and non-briefed group in the scoring of the "trunk twisting / sideways bending" risk factor for both Tasks 2 ($p < 0.05$) and Task 4 ($p < 0.01$). The briefed group observed a video clip explaining the risk categories of trunk twisting / sideways bending" while the non-briefed group could only rely on a written paragraph description. Inserting figures into the MAC to illustrate the three posture classifications may reduce the difference between the briefed and non-briefed group.



Task 3

Trunk twisting / sideways bending not assessed in carrying operations.

Figure 15 Inter- and intra-group agreement for the "trunk twisting / sideways bending" risk factor scores

4.4.2 Usability

Of the briefed group, 76% rated the usability of the "trunk twisting / sideways bending" risk factor favourably as either 1 or 2 on a 5-point scale (Figure 16). This was significantly better ($p < 0.05$) than the non-briefed group, where only 51% rated the usability favourably. This may reflect a need for illustrated examples of twisting, sideways bending and a combination of the two within the MAC. Three users from the non-briefed group wrote that it was difficult to assess twisting and sideways bending from the video. This may be partly due to the lack of

stereoscopic information, but may also be due to the lack of illustration to guide the assessment. Some users felt that "trunk twisting / sideways bending" was also a function of individual technique and wanted guidance on assessing the different techniques used by manual handlers.

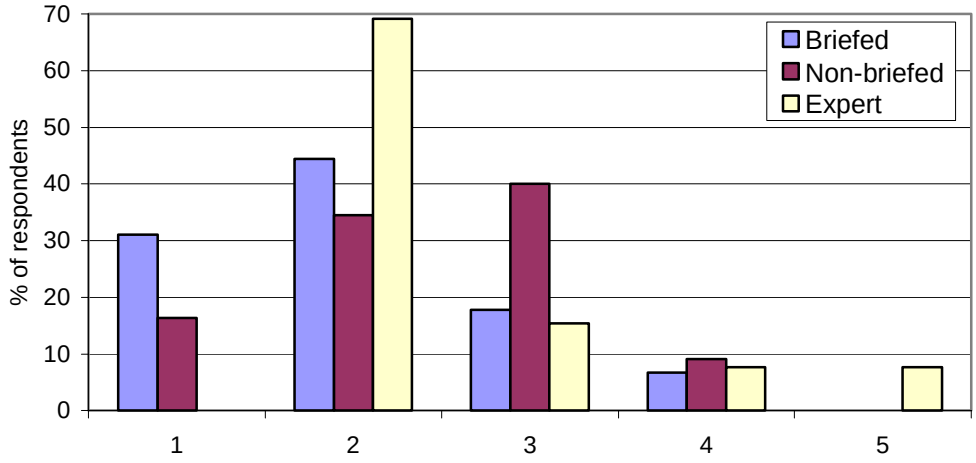


Figure 16 Usability ratings of the "trunk twisting / sideways bending" risk factor (1 = Very easy, 5 = Very difficult)

4.5 POSTURAL CONSTRAINTS

4.5.1 Levels of scoring agreement

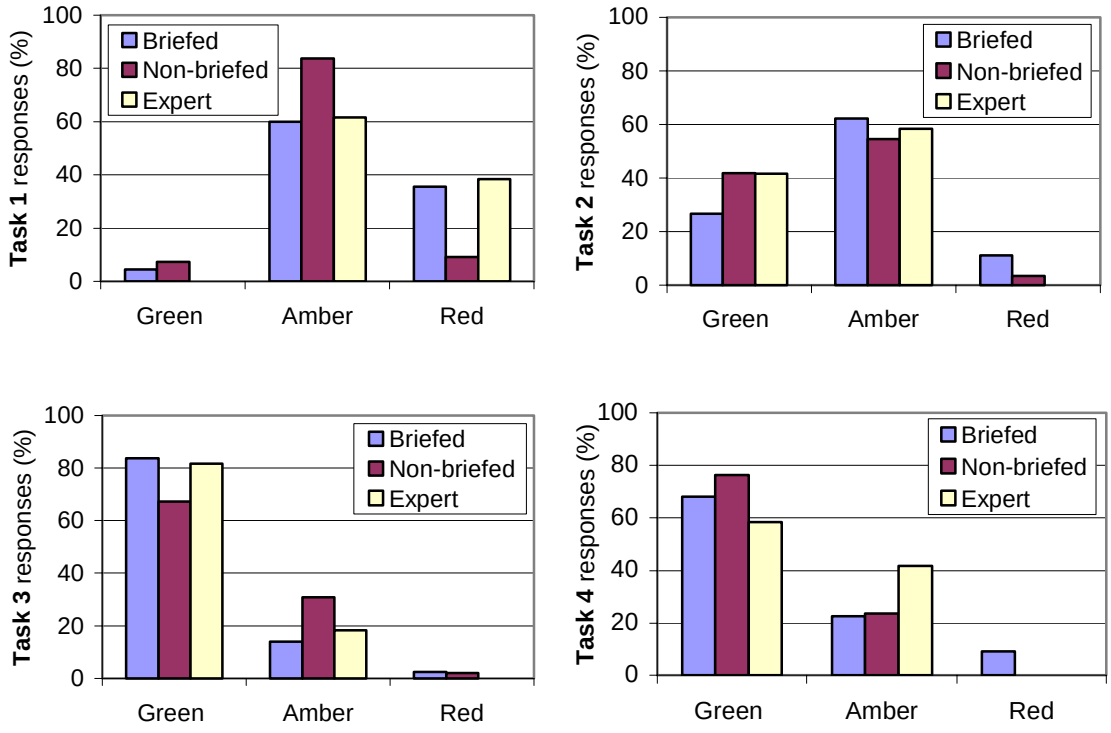


Figure 17 Inter-and intra-group agreement for the "postural constraints" risk factor scores

There was a moderate level of scoring agreement for the "postural constraints" risk factor with 60 – 80% agreement for the briefed and non-briefed groups (Figure 17). This was seen as fairly good since there is a significant amount of subjective interpretation required to make an assessment of this risk factor. It is unknown why there was a significant difference ($p < 0.01$) between the briefed and non-briefed groups for Task 1, as the "postural constraints" risk factor was not addressed in the briefing. Confounding factors such as differences between the two groups in experience or skill may be responsible.

4.5.2 Usability

The usability of the "postural constraints" risk factor was rated quite favourably with 78% of the briefed group, 60% of the non-briefed group and 54% of the expert group scoring 1 or 2 out of a 5-point scale (Figure 18). There were no significant differences between the groups, which is consistent with the fact that this risk factor was not addressed by the briefing.

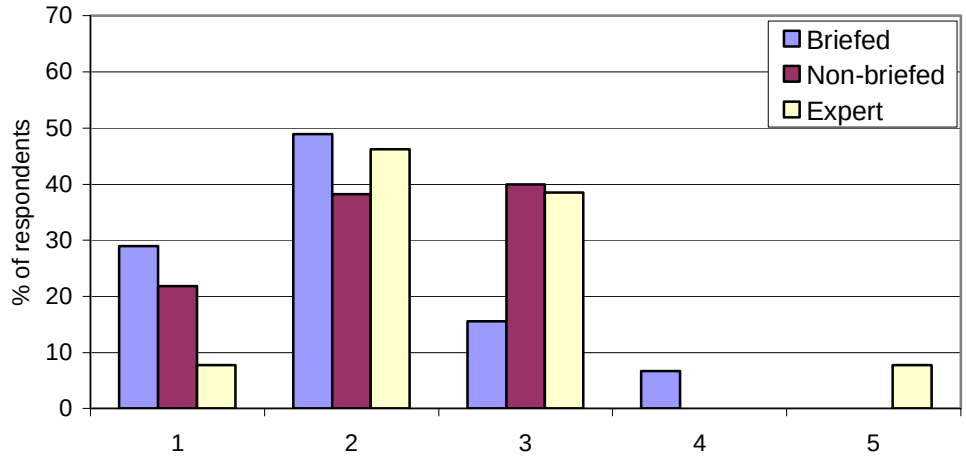


Figure 18 Usability ratings of the "postural constraints" risk factor (1 = Very easy, 5 = Very difficult)

Qualitative results suggested that further information within the descriptive paragraphs would be helpful, including the layout of information in a tabular format rather than the current written paragraph structure. Participants expressed a desire for more examples of "postural constraints" to guide their assessments, and indicated that a more intuitive name for this risk factor would be helpful.

4.6 GRIP ON LOAD

4.6.1 Level of scoring agreement

Figure 19 shows that levels of scoring agreement for "grip on load" varied between 40 – 80% depending upon how closely the handles in the videos matched the descriptions provided in the MAC. Approximately 75% of participants could recognise the bulky characteristics of the compost bales in Task 2. However, disagreement occurred when handle design or loose parts of the load had to be assessed subjectively for suitability. There were significant differences between the briefed and non-briefed groups ($p < 0.01$) and between the expert and-non expert groups ($p < 0.05$) in the assessment of the grip on the creel lips (Task 1). During the briefing, participants witnessed an example of a container with well-designed handles, which may have guided their assessment towards a higher level of risk, more congruent with expert opinion. Examples of well designed and poor handles, handholds and loose parts either within the MAC

or on the website may improve the level of scoring agreement for the "grip on load" risk factor. The significant difference between experts and non-experts, may be due to the latter having less experience of assessing tool handle design for human hand compatibility.

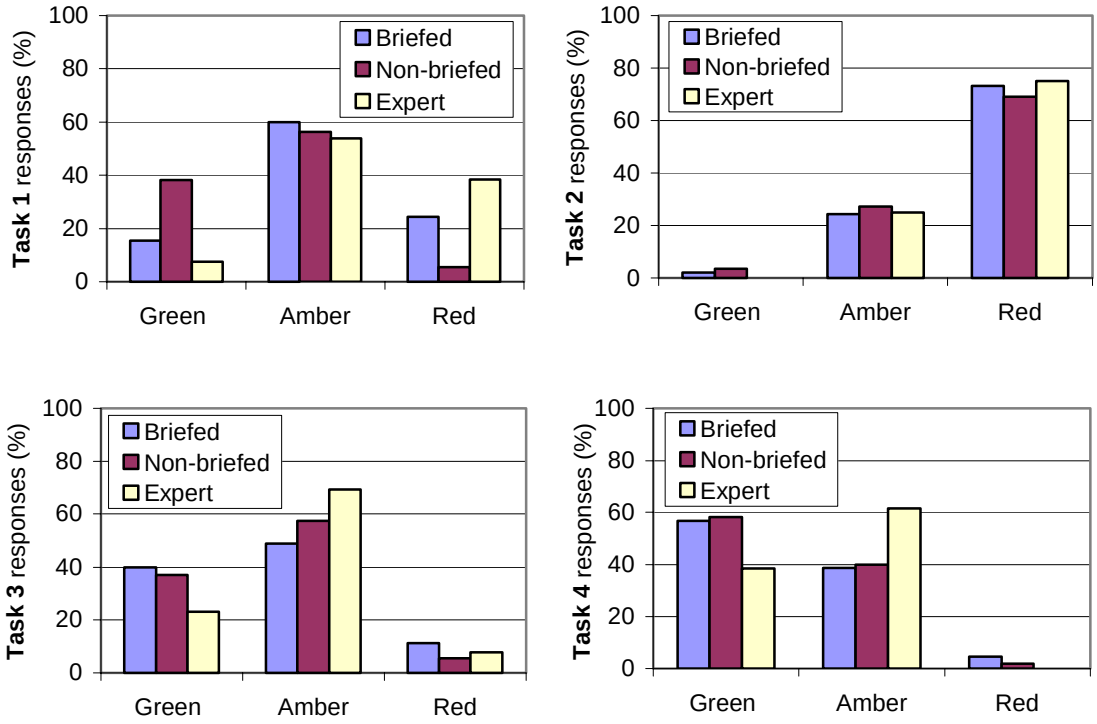


Figure 19 Inter- and intra-group agreement for the "grip on load" risk factor scores

4.6.2 Usability

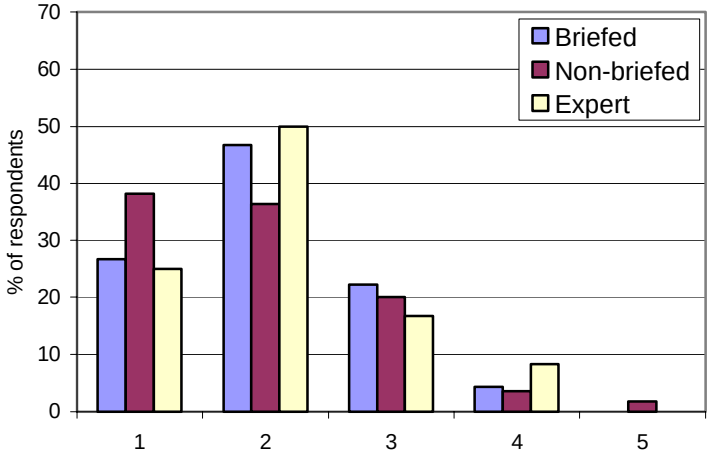


Figure 20 Usability ratings of the "grip on load" risk factor (1 = Very easy, 5 = Very difficult)

Despite some variation, 73%, 75% and 75% respectively of the briefed, non-briefed and expert groups rated the usability of the "grip on load" risk factor favourably, scoring 1 or 2 on a 5-point scale (Figure 20). Qualitative feedback indicated that problems with this factor occurred

because the grip-handle interfaces were small, and often obscured by the camera angle. In the field an assessment of the "grip on load" should be easier to perform and be more accurate than from video. Some participants also asked for guidance on assessing loads when the objects being handled are variable or where multiple grips are possible during a lift. An example given was baggage handling where loads range from square boxes to pushbikes.

4.7 FLOOR SURFACE

4.7.1 Levels of scoring agreement

Levels of scoring agreement for the "floor surface" risk factor were fairly good, with 60 – 98% agreement among the briefed and non-briefed groups (Figure 21). For Task 1 and Task 3, 100% of experts actually agreed on the "floor surface" level of risk. The level of agreement for Task 3 may be artificially high as researchers told users that the flooring was wet and contaminated since it could not be assessed visually from video. It is therefore surprising that a significant difference between the briefed and non-briefed groups ($p < 0.05$) was found for Task 3 and this must be attributed to a significant proportion of the briefed group not hearing the researchers during the assessment. It is possible that levels of agreement will improve in the field, when users can physically inspect the quality of the floor surface.

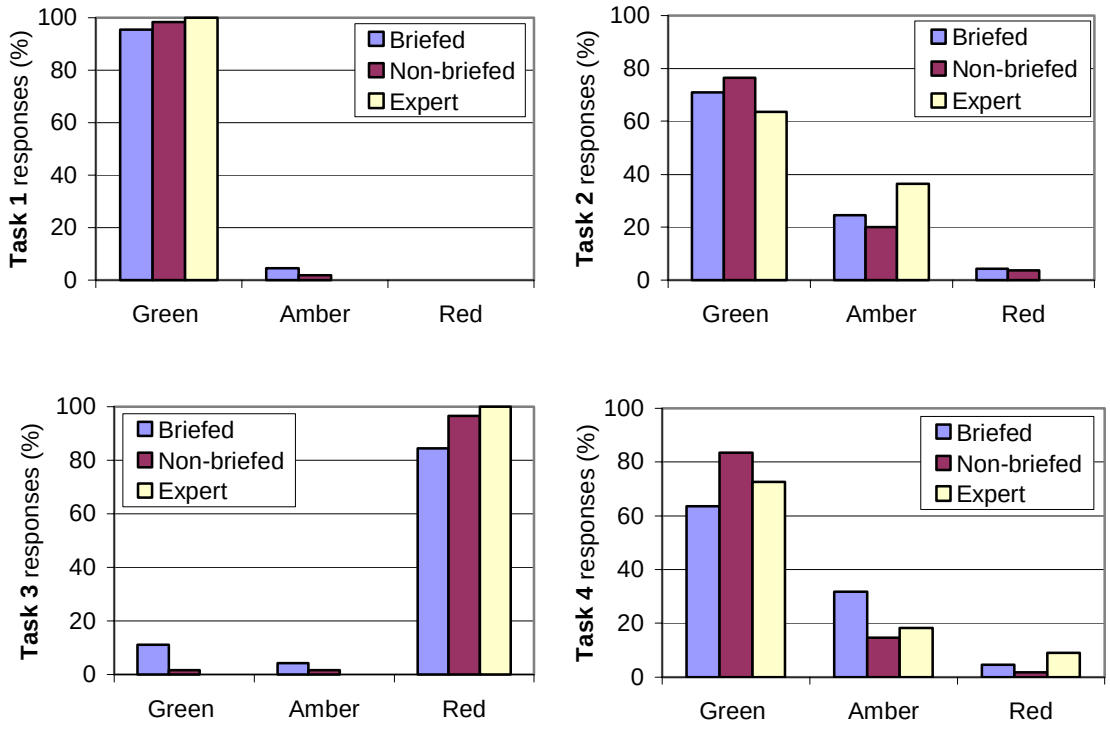


Figure 21 Inter- and intra-group agreement for the "floor surface" risk factor scores

4.7.2 Usability

Figure 22 shows that 82% of the briefed group, 75% of the non-briefed group and 83% of the expert group rated the usability of the "floor surface" risk factor favourably (1 or 2 on a 5-point scale). These values are high given that many participants felt the usability of this risk factor suffered due to the difficulty of assessing the floor condition on video. Qualitative feedback

suggested that the effect of the external environment should also be mentioned as outside floor surfaces may be in good condition in dry weather, yet wet and slippery in rain.

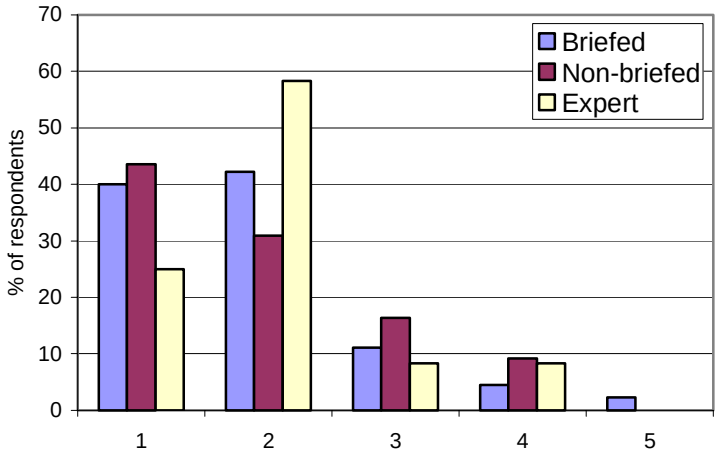


Figure 22 Usability ratings of the "floor surface" risk factor (1 = Very easy, 5 = Very difficult)

4.8 OTHER ENVIRONMENTAL FACTORS

4.8.1 Levels of scoring agreement

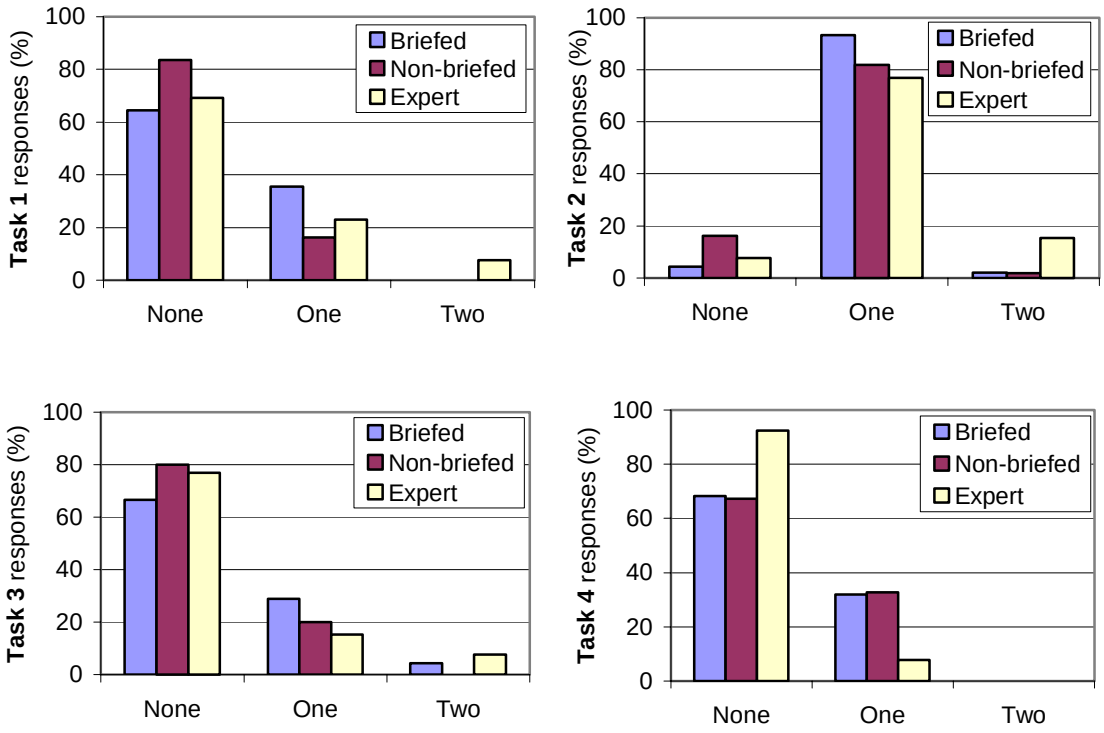


Figure 23 Inter- and intra-group agreement on the numbers of "other environmental factors" for each task

Figure 23 shows that levels of scoring agreement ranged from 64 – 98%. However, as extreme temperatures and other environmental factors are difficult to assess by video, the researchers

had to provide users with the information. Thus, caution must be exercised when extrapolating the results of this study to a field scenario. Disagreement occurred because some users chose to assess lighting conditions with the video and some users may not have heard the researchers' instructions. This would account for the significant difference ($p < 0.05$) between the briefed and non-briefed groups in Task 1. This difference was unexpected since the "other environmental factors" risk factor was not addressed by the briefing. The poor usability rating for this risk factor and the subsequent confusion in scoring also would have reduced the level of scoring agreement.

4.8.2 Usability

Figure 24 shows that only 44% of the briefed group, 74% of the non-briefed group and 58% of the expert group rated the usability of the "other environmental factors" risk factor favourably with a score of 1 or 2 on a 5-point scale.

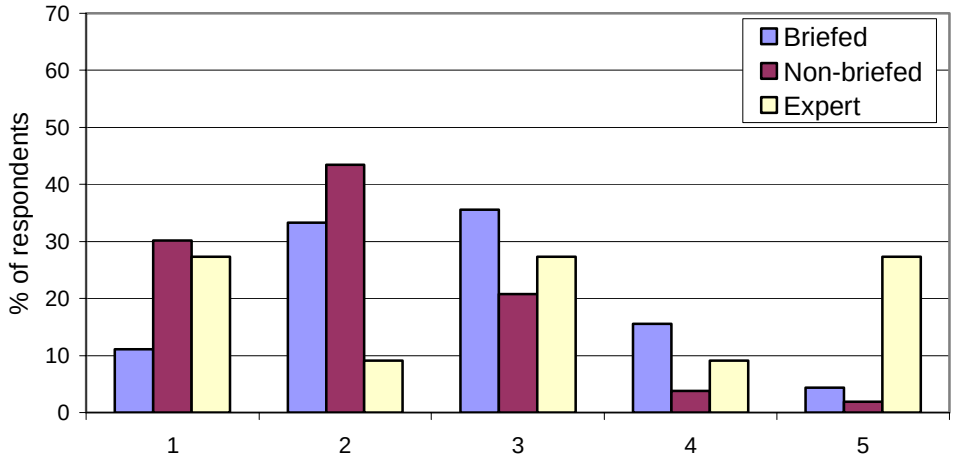


Figure 24 Usability ratings of the "other environmental factors" risk factor (1 = Very easy, 5 = Very difficult)

As the briefing did not cover this risk factor, the significant difference ($p < 0.01$) in the usability rating between the briefed and non-briefed group was unexpected. It may be that the briefed group became more accustomed to additional instruction for each risk factor provided by the briefing, and thus they might have had higher expectations regarding the clarity of those MAC risk factors not addressed by the briefing.

Qualitative feedback revealed several sources of confusion with this risk factor:

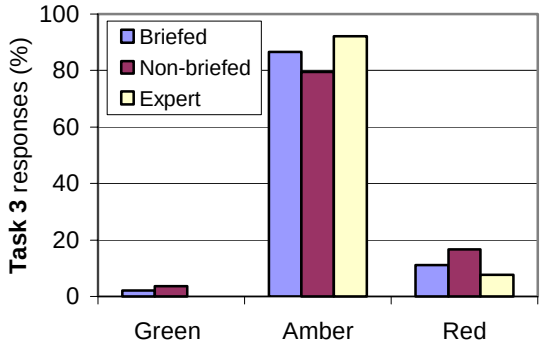
- Departure from the familiar traffic light pattern. Three amber boxes were used instead.
- There did not appear to be a "0" option
- Lack of clarity whether the scores were cumulative, or if 1 was the maximum score
- The definitions and thresholds seemed too vague

Users suggested that additional information be provided in the descriptive paragraphs and that a consistent traffic-light pattern approach be adopted. The qualitative feedback also indicated that some users wished for guidance on how to assess environmental factors that may change over time.

4.9 CARRY DISTANCE

4.9.1 Levels of scoring agreement

Figure 25 shows that the level of scoring agreement for the "carry distance" risk factor was good, ranging from 80 – 92% among the briefed, non-briefed and expert group. The level of agreement should improve within the field, as many participants mentioned that they preferred to measure the carry distance rather than estimate the distance from the video.



Task 1, Task 2 & Task 4

Carry distance not assessed in lifting or team handling operations

Figure 25 Inter- and intra-group agreement for the "carry distance" risk factor score of Task 3

4.9.2 Usability

Figure 26 shows that 80% of the briefed group, 79% of the non-briefed group and 46% of the expert group rated the usability of the "carry distance" risk factor favourably, indicating a score of 1 or 2 on a 5-point scale. The significantly ($p < 0.01$) poorer usability rating of the expert group was possibly due to the experts' common practice of measuring the carry distance with a tape measure. Some users again wished for guidance on how to assess carry distances that were highly variable, for example, in the home delivery of parcels, furniture or appliances.

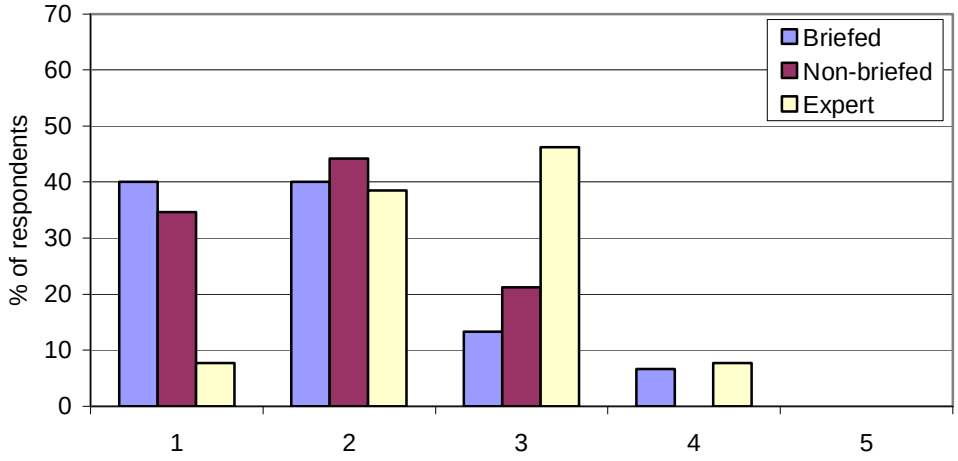
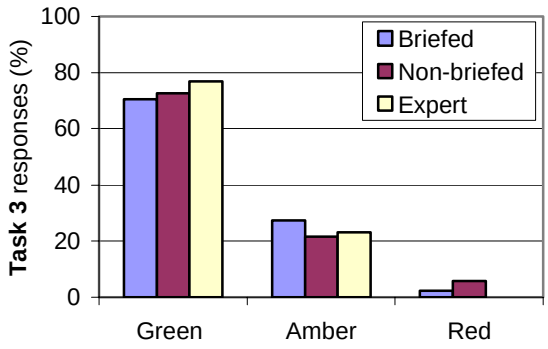


Figure 26 Usability ratings of the "carry distance" risk factor (1 = Very easy, 5 = Very difficult)

4.10 ASYMMETRICAL TRUNK / LOAD

4.10.1 Levels of scoring agreement

The level of scoring agreement for the "asymmetrical trunk / load" risk factor was fairly good, ranging from 71 – 77% for the briefed, non-briefed and expert group (Figure 27). As this risk factor was only assessed in Task 3 and involved a simple symmetrical load, it is unknown whether a similar level of agreement could be obtained for more asymmetrical loads. For example, if an operator carried an object in each hand to the side of their body, could users reliably differentiate between the amber and red levels of risk? Further testing is recommended for this risk factor. The current variability was mostly due to users in the food industry who, unsure whether the container contained solid or liquid ingredients, assumed a 'worst-case scenario' and scored amber for an asymmetrical load of liquid ingredients.



Task 1, Task 2 & Task 4

Asymmetrical trunk / load not assessed in lifting or team handling operations

Figure 27 Inter- and intra-group agreement for the "asymmetrical trunk / load" risk factor score of Task 3

4.10.2 Usability

Figure 28 shows that 84% of the briefed group, 75% of the non-briefed group and 31% of the expert group rated the usability of the "asymmetrical trunk / load" risk factor favourably, scoring a 1 or 2 on a 5-point scale.

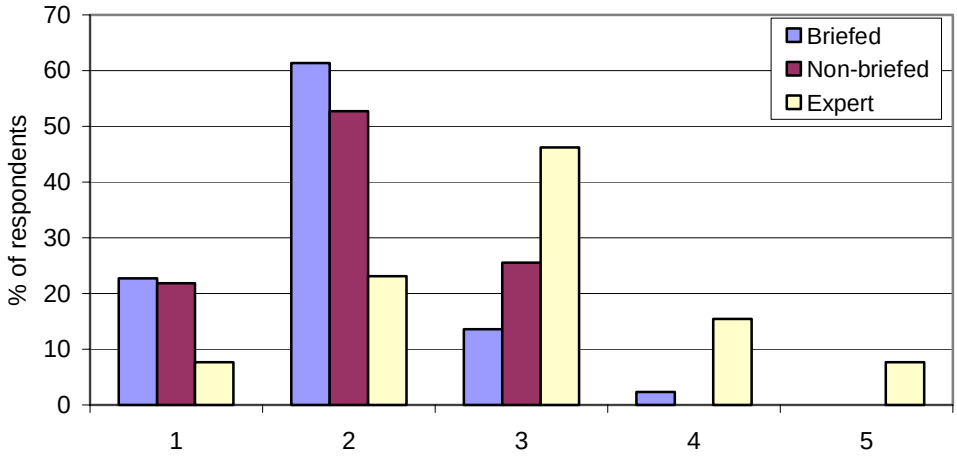


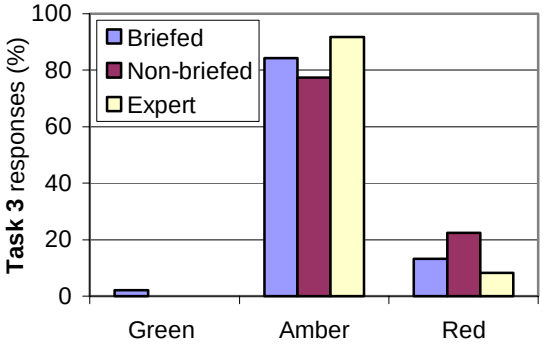
Figure 28 Usability ratings of the "asymmetrical trunk / load" risk factor (1 = Very easy, 5 = Very difficult)

There was a significant difference between experts and non-experts in the usability rating of this risk factor ($p < 0.01$) but the experts did not explain why this risk factor was less easy to use. Some users indicated that further description for this risk factor should be provided. During the assessment of Task 3, users noticed that the MAC provided little guidance for scoring shifting or unpredictable loads, such as a tub filled with liquid.

4.11 OBSTACLES EN ROUTE

4.11.1 Levels of scoring agreement

The level of scoring agreement for the "obstacles en route" risk factor was good, ranging from 77 – 92% among the briefed, non-briefed and expert groups (Figure 29). The slight variation occurred as the steep slope of the steps led some users to assess the steps as a ladder instead. Some users felt strongly that the steepness of the steps created a higher level of risk than suggested by the MAC and thus scored a red level of risk rather than amber.



Task 1, Task 2 & Task 4

Obstacles en route not assessed in lifting or team handling operations

Figure 29 Inter- and intra-group agreement for the "obstacles en route" risk factor score of Task 3

4.11.2 Usability

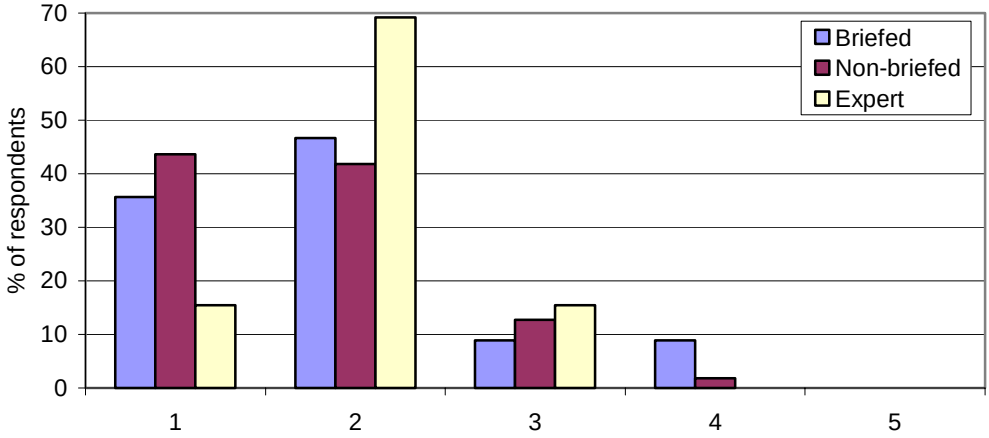


Figure 30 Usability ratings of the "obstacles en route" risk (1 = Very easy, 5 = Very difficult)

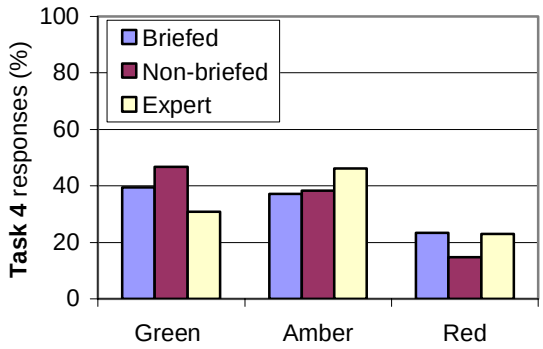
Figure 30 shows that 82% of the briefed group, 85% of the non-briefed group and 85% of the expert group rated the usability of the "obstacles en route" risk factor favourably, scoring 1 or 2

on a 5-point scale. The primary usability problem mentioned in the qualitative data was confusion over the interpretation of steps and ladders. One user raised a concern regarding how to assess "obstacles en route" if they often vary over time. Task 3 only required the assessment of one obstacle; however, the researchers anticipate possible usability concerns if more than one obstacle is present. In particular, should a route with a steep slope and steps be scored as A/2 (non-cumulative), A/4 (cumulative score) or R/4 (cumulative colour and score)? Although the descriptive paragraphs instruct users to 'total the score', it may be inappropriate to associate an amber level of risk with a greater numerical score than a red level of risk.

4.12 COMMUNICATION & COORDINATION

4.12.1 Levels of scoring agreement

Figure 31 shows that the level of scoring agreement for the "communication and co-ordination" risk factor is poor, ranging from 40 – 47% among the briefed, non-briefed and expert groups. As there was no audio feedback on the video of Task 4, users could not assess auditory communication amongst the members of the team handling operation. Rather, users were required to assess this risk factor using non-auditory cues, visually assessing the co-ordination of the operation only.



Task 1, Task 2 & Task 3

Communication and co-ordination are not assessed in lifting and carrying operations

Figure 31 Inter- and intra-group agreement for the "communication and co-ordination" risk factor of Task 4

This situation explains much of the variation witnessed, as does the highly subjective interpretation required to assess this risk factor. It is probable that users could assess "communication and co-ordination" with greater reliability in the field.

4.12.2 Usability

Figure 32 shows that 58% of the briefed group, 53% of the non-briefed group and 23% of the expert group rated the usability of the "communication and co-ordination" risk factor favourably, scoring a 1 or 2 on a 5-point scale. The significantly poorer ($p < 0.01$) usability rating reflects the difficulty of assessing communication from a video lacking audio feedback. Although this risk factor is expected to be more usable in the field, qualitative feedback highlighted several possible improvements. Some users found the information in the descriptive paragraphs too limited to guide their subjective assessment of 'good', 'reasonable' or 'poor' communication and co-ordination. This was especially pertinent for non-verbal cues, which users had to rely on in this study. In addition, numerical scores were not provided within the descriptive paragraphs for this risk factor. This was inconsistent with all other MAC risk factors and required users to rely on the flowchart to find the corresponding level of risk. With

little experience performing assessments with the MAC, users relied heavily upon the descriptive paragraphs and score sheet, paying less attention to the flowcharts.

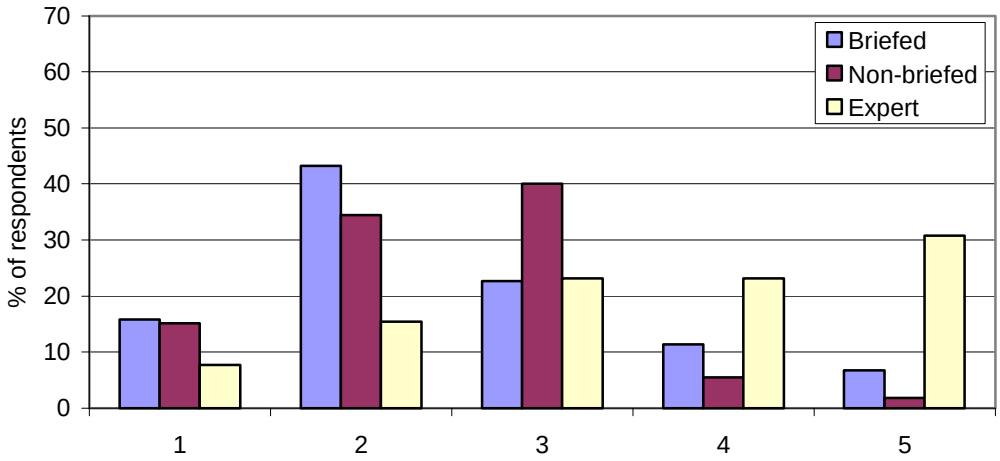


Figure 32 Usability ratings of the "communication and co-ordination" risk factor (1 = Very easy, 5 = Very difficult)

4.13 SUMMARY OF THE DETAILED RISK FACTOR ANALYSIS

For most MAC risk factors, a majority of users rated the usability favourably. This allowed a fairly reliable assessment of the tasks among the briefed and non-briefed groups. A larger sample size of experts would be preferable to accurately assess expert reliability and usability. Some occasions of lower scoring agreement, poorer usability ratings or significant differences between groups were evident. However, where these problems occurred, they could usually be explained as either a limitation of the video-based study or rectified with information obtained from verbal and written qualitative feedback.

5 DISCUSSION OF QUALITATIVE FEEDBACK

5.1 PERCEPTIONS OF SUITABILITY

Many users recognised that the deliberate simplification of the MAC made it unsuitable for assessing some manual handling operations. Based on the available evidence, however, it seems reasonable to assume that very few users would have the knowledge and experience to recognise all the limitations of the MAC. It was therefore realised that it would be prudent to state clearly the limitations within the MAC itself and provide more detail, via for example the website, to minimise the problem of misapplication.

According to users' feedback, it was felt that the MAC would not be appropriate for assessing these general activities:

- Pushing and pulling
- Rolling
- Jobs with large variations in loads, frequencies or activities
- Animate loads
- Tasks involving both lifting and carrying
- Tasks involving both team lifting and carrying
- One handed lifts

The MAC appears to be best suited for well-defined tasks or jobs with relatively constant risk factors. Members of the parcel carrier industry were the most vocal on the problems this raised for them, because of the extreme variability in the weight of parcels that their organisations handle. For example, a parcel distribution centre may handle 14000 individual packages per night with weights ranging from a few grams to 50 kg. The MAC does not state clearly whether the load weight value should be the mean, mode, maximum or something else, and how to practically determine this value. Concerns were also raised with respect to assessing frequency of handling when it is not fixed but variable, or the hand distance from the lower back, the presence of trunk twisting, the quality of the grip on the load, the carrying distance, the presence of the lsited environmental factors, and the existence of obstacles en route.

The issue of variability in frequency of handling is important issue, given that relatively few jobs place constant, homogenous physical demands on workers. However, the scientific literature currently tells us little about the relative effects of differing work patterns on worker health. For example, the psychophysical studies underlying the load/frequency charts (Snook and Ciriello, 1991) report maximum acceptable weights for lifting at constant rates over a working day. While it has been suggested that the frequency of handling which occurs in bursts of less than 15 minutes duration should be averaged over the length of the task cycle including subsequent recovery periods (Waters *et al.*, 1994), it appears that no experimental or epidemiological studies have yet addressed this issue.

The second point to consider is that it is the very nature of tools such as the MAC to trade simplicity for versatility. That is to say, that an attempt to account for every possible influence on manual handling risk would necessarily require a highly detailed and cumbersome tool that

would not meet the criterion that the MAC should be quick and easy to use. A conscious decision has been made not to quantify the effects of variability within risk factors such as load weight and frequency.

To standardise the study and to err on the side of caution as good health and safety practice, the researchers instructed users to assess the 'worst case scenario'. However, no such guidance was written within the version of the MAC tested, nor do the researchers believe that this is always the best approach, for example in the case of the load weight. However, for the "hand distance from lower back" and "vertical lift region" risk factors, the updated version of the MAC- now states: "Always assess the worst case scenario".

That many users felt the MAC would not be suitable for assessing manual handling operations with both lifting and carrying or team handling and carrying may seem surprising. The MAC states: "Where the task involves both lifting and carrying, both should be assessed". The response may reflect disapproval of having to perform a separate lifting and carrying assessment if both activities are present in a single manual handling operation.

Users also described several specific manual handling activities for which the MAC would not be a suitable assessment tool. These included:

- Taking large bars out of drums
- Mould handling
- Hand-paddling mixes (mixing materials in bins)
- Courier deliveries to customer sites
- Heavy oven band lifting
- Heavy object lifting (individual)
- Tailors press from trolley to platform
- Handling flexible pipes (e.g. petrol/diesel tank drivers)

For the most part, explanations for these specific exclusions will be similar to those in the discussion of the general activities for which the MAC is unsuitable, above. For example, assessing courier deliveries with the MAC is problematic due to the variable nature of the tasks. Heavy oven band lifting, however, is clearly within the intended scope of the MAC. It is possible that the user misunderstood either the question or the applicability of the MAC.

5.2 GUIDANCE

5.2.1 Guidance on the application of the MAC

A recurrent theme during the study was the request for more explicit description on the general nature, purpose and status of the MAC. Some users requested more description on the purpose of applying the MAC. However, most comments were with respect to the legal standing of the MAC and its relationship to the Manual Handling Operations Regulations 1992 (HSE, 1998). Many users realised that some manual handling risk factors that should be considered in a full risk assessment were omitted from the MAC. Some users also highlighted limitations of the MAC, such as its 'inappropriateness' to assess manual handling of animate loads. They felt strongly that these limitations should be presented clearly to prevent misapplication of the

MAC. The inside front cover of the MAC has been substantially rewritten to address as many of these concerns as possible within the space limitations.

There were some general and specific criticisms regarding a lack of clear instruction on how to carry out an assessment with the MAC. As a result a section has been included in the Introduction on "How to complete a MAC assessment" which gives step by step instructions on how to proceed through the tool using the flowcharts, the assessment guides and the score sheet.

Many expert users felt that, in line with HSE's policy of adopting an ergonomics approach to manual handling assessment, the MAC should actively foster employee consultation during the risk assessment process. Although worker input was not actively incorporated into the scoring process, design features of the MAC, such as its traffic light colour scheme and flowchart structure are familiar to workers and have proven to engage line managers as well as health and safety representatives. In addition, there is potential to promote the MAC with posters and thereby increase the awareness of manual handling risk factors at the workplace. As a reminder, the instructions on "How to complete a MAC assessment" now state: "Consult employees and safety representatives during the assessment process. Where several people do the same task, make sure you have some insight into the demands of the job from all workers' perspective."

5.2.2 Guidance on how to interpret the results of a MAC assessment

Many users were initially confused that the final score did not correspond to a colour indicating the overall level of risk. They mentioned that this feature would be extremely beneficial when trying to justify health and safety expenditures to senior management. The difficulties of systems which seek to combine risk scores from a wide range of risk factors have been discussed by Pinder (2002) in the context of benchmarking the MAC against other manual handling assessment tool. There is currently a lack of scientific evidence to support the assignment of specific numeric scores to combined risk levels, and this is unlikely to change without further targeted research.

The verbal guidance provided in the sessions was that there are two potential uses for the risk scores. One was as a rough inter-job prioritisation method for taking action on a wide range of jobs. Although a score of 23 is not necessarily a higher risk than a score of 22, starting with the scores of 28 and progressing systematically towards the scores of 15 would be a reasonable approach, all else being equal. The other use was as an intra-job comparison. The final score could be calculated before and after a change to an operation to see if a reduced level of risk actually occurred. Alternatively, this could assist health and safety representatives compare several potential job modifications before deciding upon the risk control(s) to implement. The MAC now states that: "The total scores help prioritise those tasks that need most urgent attention and help check the effectiveness of those improvements." More detailed guidance is also given in the "Frequently Asked Questions" page of the website.

Expert users were divided with regards to whether the MAC should be used for seeking solutions. Concerns were raised that some users may form conclusions about solving manual handling problems based on an incomplete set of facts, potentially creating new problems. However, most users found that the individual colour bands allowed them to see potential solutions that they otherwise might not have considered. It was thought that both views are valid depending upon the specific application of the MAC. This conflict lent further support to the need to link the MAC explicitly to the more comprehensive approach of risk assessment in the HSE guidance on the 1992 Manual Handling Operations Regulations (HSE, 1998).

5.3 FORMATTING

The overwhelming majority of users were very positive about the MAC as a whole. Many of the positive comments referred to the formatting concept of the MAC and included: the colour scheme; the traffic light pattern of risk classification; the step-wise approach; and the graphic illustrations. Some users also envisioned potential for electronic versions of the MAC in the forms of an interactive web version, downloadable executable file and a Personal Digital Assistant (PDA) version.

The four primary formatting concerns involved: the colour scheme, the design of the score sheet; the order of the score sheet with respect to the flowcharts; and the separation of the MAC into three assessment tools (lifting, carrying and team handling).

5.3.1 Colour scheme

Many users spoke positively of the green / amber / red traffic light colour scheme. It was thought though that similarity between the red and purple colours could make it difficult for some users to differentiate between high and very high levels of risk for the "load weight / frequency" risk factor. Also, the yellow shading of the risk factor boxes on the flowcharts was felt by some to be too similar to the amber risk colour.

Some users questioned the usability of the MAC for people with defects in their colour vision (approximately 8% of the male population and 0.5% of the female population). The MAC incorporates redundant textual labels to enable full use even in greyscale. However, in the interest of usability, further research for users with defective colour vision, in particular those with reduced spectral colour discrimination for green, yellow, orange and red, would provide greater reassurance.

5.3.2 Score sheet format

It became apparent that the format of the score sheet did not guide users to perform a separate assessment for lifting, carrying or team handling operations (Figure 33). It was observed that , rather than follow the flowchart order, many users would work down the score sheet and use the descriptive paragraphs and flowcharts solely to assess the level of risk. This occurred because after users entered the colour band and numerical score for one risk factor on the score sheet, it was intuitive to look to the row below on the score sheet to determine which risk factor to assess subsequently. Consequently, it was found that 56% of all users had scored additional risk factors not prescribed by the flowcharts (e.g., a carry distance during the Task 1 creel lifting assessment).

Another concern was that the order of the risk factors on the score sheet did not correspond to the order of the risk factors on the carrying and team handling flowcharts. Thus, users were observed on several occasions to insert risk factors into the wrong box on the score sheet. For example, it was possible for users to insert a "carrying distance" score, as the second risk factor on the carrying flowchart, into the score sheet box for the "hand distance from the lower back". Although users were instructed to avoid this error, this usability problem may have been a major confounder for some of the reliability results, in particular the levels of agreement for Task 3. Rearranging the order of the risk factors on the flowcharts to match the order on the score sheet could reduce this confusion, although at the sacrifice of listing the risk factors in a priority order.

Insert the colour band and numerical score for each of the risk factors in the appropriate boxes below, with reference to your assessment using the tool		
Risk Factors	**Colour Band (G, A, R or P)**	**Numerical Score**
Load weight and lift/carry frequency		
Hand distance from the lower back		
Vertical lift region		
Trunk twisting / sideways bending Asymmetrical trunk / load (carrying)		
Postural constraints		
Grip on load		
Floor surface		
Other environmental factors		
Carry distance		
Obstacles en route (carrying only)		
Communication and co-ordination (team handling only)		
TOTAL SCORE *(see page 13)**		
Other risk factors e.g. individual factors, psychosocial factors, etc.		

Figure 33 Tested MAC score sheet

Figure 34 shows a revised score sheet with two major modifications:

- The colour band and numerical score columns are split into three separate columns for lifting, carrying and team handling. This should guide users to perform a separate assessment for lifting, carrying and team handling operations. Where a task involves lifting and carrying, it allows scoring of each component on a single score sheet, which should reduce the paperwork of a MAC assessment.

- For each lifting, carrying and team handling operation, risk factors that do not appear on the flowcharts are shaded on the corresponding column of the score sheet. This should remind users that only the risk factors on the flowcharts need to be scored.

Insert the colour band and numerical score for each of the risk factors in the appropriate boxes below, with reference to your assessment using the tool						
Risk Factors	**Colour Band (G, A, R, or P)**			**Numerical Score**		
	Lift	Carry	Team	Lift	Carry	Team
Load weight and lift/carry frequency						
Hand distance from the lower back						
Vertical lift region						
Trunk twisting / sideways bending Asymmetrical trunk / load (carrying)						
Postural constraints						
Grip on load						
Floor surface						
Other environmental factors						
Carry distance (carrying only)						
Obstacles en route (carrying only)						
Communication and co-ordination (team handling only)						
Other risk factors e.g. individual factors, psychosocial factors, etc	**TOTAL SCORE:**					

Figure 34 Revised MAC score sheet

5.3.3 Separation of lifting, carrying and team handling operations

Many users emphatically expressed frustration with the three separate assessments for lifting, carrying and team handling operations. Users felt that producing two or three separate assessments for the same task would be too onerous, and be seen by managers and employees in their organisation as an unnecessary "paper-pushing exercise". In addition, "real world" jobs are rarely divided into just lifting, carrying or team handling. For instance:

- Carrying operations typically involve an initial lift or "vertical lift region"
- Loads that are lifted can be asymmetrical, yet this risk factor is only considered in the carrying assessment.

- Team handling operations most often involve a “carry distance” and “obstacles en route”.

The division of the MAC into three separate charts was often not intuitive. Some users from both the briefed and non-briefed groups were observed to apply the carrying flowchart for an assessment of a lifting operation and vice versa. During the session, the researchers had to constantly remind users to apply the relevant flowchart. There were other usability concerns that arose with the development of three separate assessment tools in the MAC:

- Several users were confused that the risk factors and letter descriptors did not correspond between the three tools (i.e. they lacked one to one mapping). For the lifting flowchart, “F” corresponded to the “grip on load”, while for the carrying flowchart, “F” corresponded to “postural constraints”.

- Several users experienced problems transferring the relevant data of three separate assessments to a single score sheet.

- Users preferred the use of tables in the descriptive paragraphs to explain each risk factor, rather than a paragraph of text. However, the current format did not provide enough space for additional tables. Limited space was a constant challenge when designing the MAC; however, substantial repetition was apparent within the MAC, especially amongst the common risk factors on each flowchart.

5.3.4 Feasibility of creating a single chart format

To address the limitations of three separate charts, the feasibility of a single chart format was explored. A draft of a possible single MAC flowchart is shown in Figure 35.

The basic structure of the draft MAC with the single flowchart compared to the version of the MAC that was tested is shown in Table 10. This layout would allow a larger two-page flow chart and a single three-page descriptive paragraph section while still reducing the MAC by four pages. It would also require fewer steps to perform an assessment with the MAC. Users would have to be cued to use the appropriate “load weight / frequency” graph, perhaps by placing them on facing pages.

Alternatively, the same single page flowchart could include a score sheet (Figure 36). This may further reduce some of the usability issues faced when flipping between pages to record scores or when linking scores back to original criteria. It would also allow an additional page for the descriptive paragraphs. However, the limitation of this approach is that there is insufficient space to capture some of the additional details that the separate score sheet prompts the user to input. Although further testing and refinement will be required to fully address the usability issues raised in this study, the balance of evidence strongly suggests a need for combining the three charts into one.

Experienced users of the MAC feel that the single page flowchart format has face validity and substantially addresses the concerns of separating a manual handling operation into a separate lifting, carrying and team handling assessment. Although no changes were made to the underlying mechanics of the tool, there would be a ripple effect of changes to various aspects of the MAC, and consequently a new set of minor “challenges” would likely arise that would need to be resolved systematically.

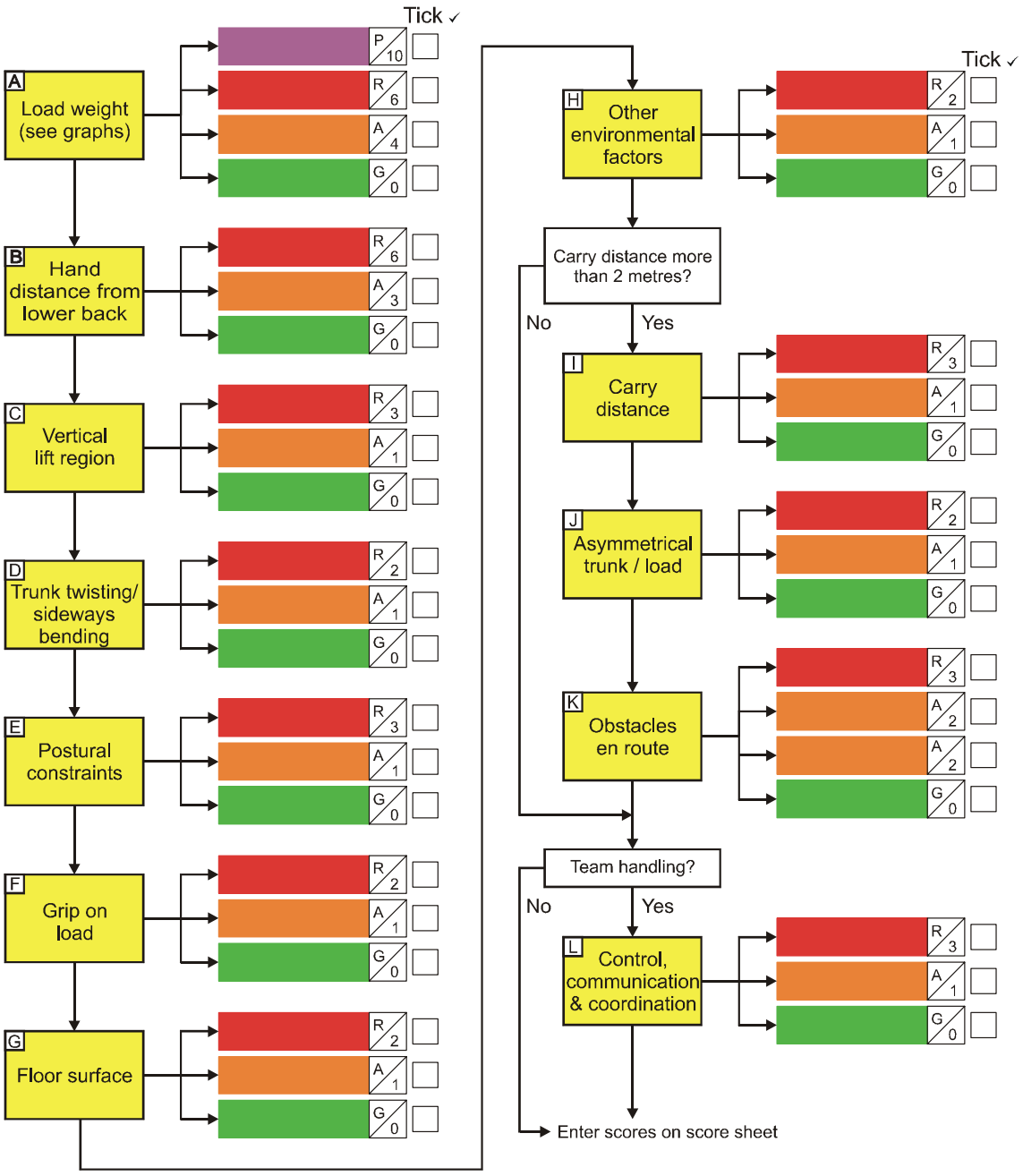
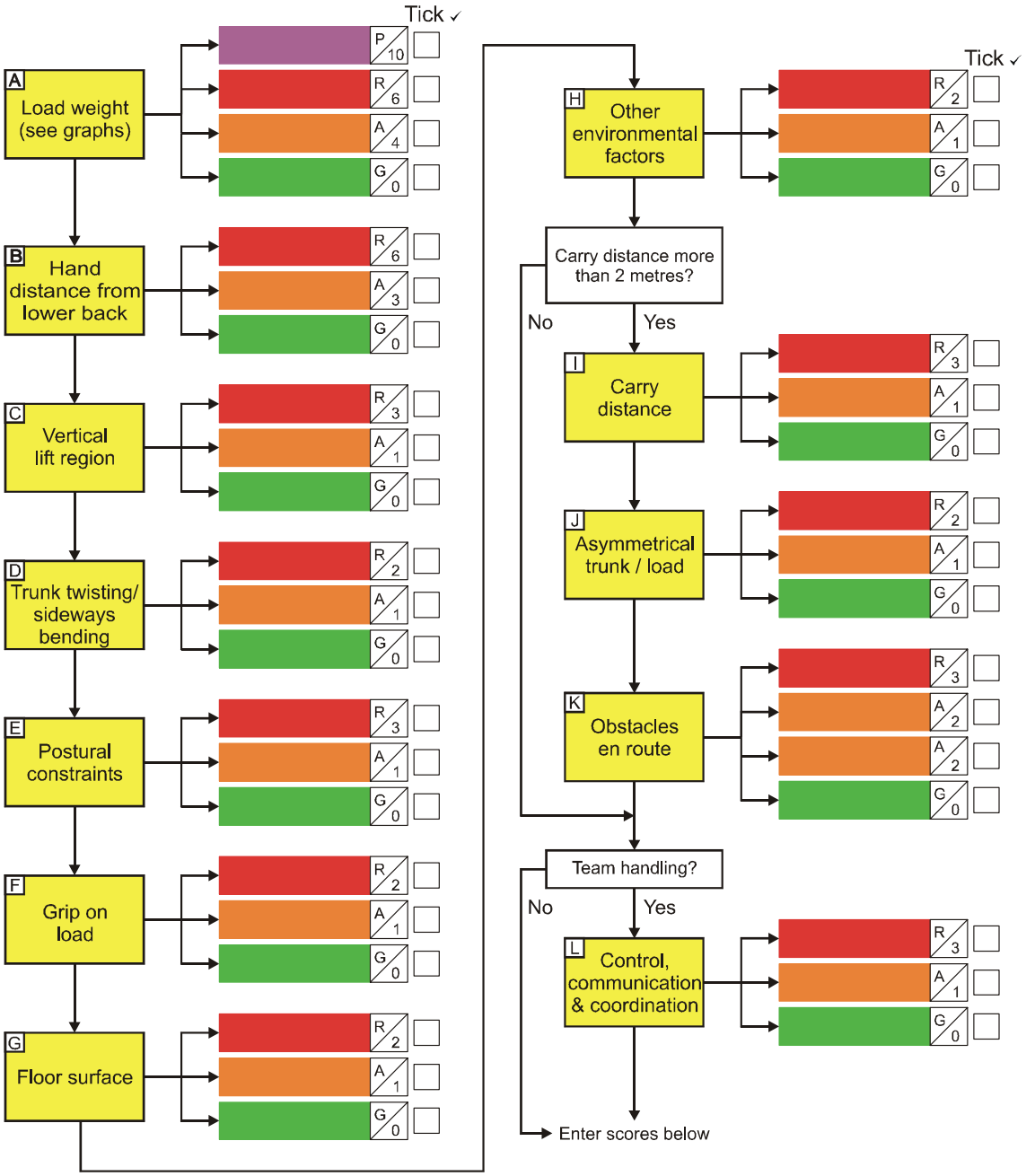


Figure 35 Proposed layout of a single page format for the MAC flowchart

Table 10 MAC page layout with single page flowchart format

Page #	*Tested layout*	*Proposed layout*
1	Front cover	Front cover
2	Inside front cover	Inside front cover
3	Contents	Contents
4	Lifting flowchart	Flowchart (facing pages)
5	Lifting load / frequency graph	Flowchart (facing pages)
6	Lifting descriptive paragraphs	Lifting load / frequency graphs (centre fold)
7	Lifting descriptive paragraphs	Carrying load / frequency graphs (centre fold)
8	Carrying flowchart	Descriptive paragraphs
9	Carrying load / frequency graph	Descriptive paragraphs
10	Carrying descriptive paragraphs	Descriptive paragraphs
11	Carrying descriptive paragraphs	Descriptive paragraph or Score sheet
12	Team handling flowchart	Risk classification
13	Team handling descriptive paragraphs	
14	Team handling descriptive paragraphs	
15	Score sheet	
16	Risk classification	



Risk Factor	A	B	C	D	E	F	G	H	I	J	K	L	Total score
Colour (P/R/A/G)													
Score													

Figure 36 Proposed layout of a single page flowchart with integrated score sheet

6 CONCLUSIONS AND RECOMMENDATIONS

An iterative process has been employed since the initial development of the MAC to continually improve and refine its structure and presentation. While early efforts were focused on developing a tool that would be suitable for inspectors (Monnington *et al.*, 2002), more recent focus has been on reliability and usability for both inspector and non-inspector use. As a result of these efforts, the underlying mechanics of the tool have remained quite stable, while the presentation and layout have undergone numerous revisions.

Although the version of the MAC field-tested with NRPs showed reasonable reliability and good usability, substantial improvements to both of these are possible and necessary. The relationship between the reliability data and the usability data was substantial. The majority of the failings in the reliability data could be explained and accounted for by specific usability issues. When taken together, the quantitative and qualitative usability data provided clear direction in most cases for specific changes to the structural, descriptive and presentational aspects of the MAC.

RECOMMENDATIONS

Thirty recommendations were made as a result of the examination of the usability and reliability of the MAC in this study. Table 11 cross-references the sections of the study which lead to each recommendation and lists the changes made as a result or the reason for deferring accepting the change. The recommendations can be summarised into four broad categories:

- Clearly define the limitations, legal standing and appropriate usage of the MAC.
- Expand the descriptive paragraphs to include additional colour coded tables, diagrams and explanatory text.
- Modify the score sheet design to minimise the need to flip between pages to record scores; to allow separate scoring of lifting, carrying and team handling operations; and to prevent users from scoring inappropriate risk factors.
- Adopt, test, and refine a single page format for the MAC flowcharts.

Changes based on seventeen of the recommendations were incorporated into the version of the MAC released to the public in August 2003 (HSE and HSL, 2003). Two of these had additional supporting documentation provided on HSE's website (http://www.hse.gov.uk/msd/) at the same time. Four were addressed soley through material made available on the website.

Acceptance of the remaining nine recommendations was deferred. The most significant ones related to the current MAC format of separate charts for lifting, carrying and team handling and could only be properly addressed by combining the three flowcharts into one. The researchers believe that there is ample evidence to support this change because such an approach would directly simplify the risk assessment process, allow the lift, carry and team handling aspects of a single task to be assessed together, and alleviate the space constraints which are the root of many of the usability issues. However, it was felt that such a change to the structure of the MAC would require further testing to address likely "ripple effects" that would require further changes to the descriptive and presentation aspects to ensure optimum usability.

Table 11 Recommendations for changes to the MAC made as a result of this study and the outcomes

No.	*Recommendation*	*Section*	*Change*	*Details of change*
1	State that the lifting chart is also applicable to lowering tasks.	4.1	MAC	The introduction has been modified to make this clear.
2	Improve guidance on addressing variability when measuring risk factors, with special emphasis on the "load weight / frequency" risk factor.	4.2 & 5.1	Defer	Unfortunately, the gaps in the scientific literature do not permit us to provide defensible answers to every question. Also, it is the nature of tools such as the MAC to trade versatility for simplicity. Users of the MAC must make informed decisions as to when it is appropriate to use the maximum, mean, or some other value when considering loads and frequency. Difficulty in determining the appropriate values to use may indicate that a more detailed risk assessment is indicated.
3	Clarify the difference between the "hand distance from lower back" and "vertical lift region" risk factors (i.e. horizontal vs. vertical reach components).	4.2	MAC	This has been clarified in the Assessment Guides for all three operations. This will discourage inclusion of the vertical component of the lift in this factor and hence discourage double counting of the vertical component.
4	Rename the "hand distance from lower back" risk factor to more accurately reflect the horizontal distance that is being assessed.	4.2	Defer	Limited space in the flow charts would make this change difficult. It is believed that the change in the assessment guides will obviate the problem.
5	Remove conflicting references to the "vertical lift region" reference points.	4.3	MAC	The words "lift from" have been removed from the descriptors accompanying the illustrations. This has removed the erroneous suggestion that only the starting point needs to be considered.
6	Provide illustrated examples of the three posture classifications for the "trunk twisting / sideways bending" risk factor.	4.4	Defer	Space limitations did not allow this change to be made. In addition, further research will be needed on the most appropriate method of visually representing the degree of twisting or bending required to indicate a positive response.
7	Provide further written description and add illustrated examples of postural constraints; poor handles, handholds, and loose parts; other environmental factors; communication and co-ordination; using a colour-coded tabular format.	4.5, 4.6, 4.8, & 5.3.3	Web	Space limitations did not allow these changes to be made in the MAC but extra information was included in the on-line guidance. It is anticipated that a single-page flowchart format would allow 2-3 additional pages to be used for enhancing the descriptive paragraphs.
8	Make reference to the impact of the external environment on the floor surface.	4.7	Web	This change has been included in the on-line guidance within the page offering general risk reduction measures, but not in the MAC itself.
9	Adopt a traffic light pattern for scoring the "other environmental factors" risk factor.	4.8	MAC	The traffic light pattern has been adopted.

10	Provide clear criteria for scoring 0, 1, 2, or 3 of the "other environmental factors".	4.8	MAC	The following definitions have been provided: Green means no factors present, and scores 0. Amber means one factor present, and scores 1. Red means two or more factors present, and scores 2.
11	Combine the two amber risk factors for the "obstacles en route" risk factor.	4.11	Defer	Further usability testing is required before changing this risk factor.
12	Remove the cumulative scoring of factors for the "obstacles en route" risk factor.	4.11	Defer	Further usability testing will be performed on this risk factor when developing future versions of the MAC.
13	Provide numeric scores in the descriptive paragraph of the "communication and co-ordination" risk factor.	4.12	Web	This change has been included in the on-line guidance, but not in the MAC.
14	Clearly state the limitations of the MAC.	5.1	MAC/ Web	The introduction now contains a highlighted section stating what the MAC is ***not*** appropriate for.
15	Describe the legal standing of the MAC and its relationship to the Manual Handling Operations Regulations 1992 (as amended).	5.2.1	MAC/ Web	The introduction now contains a section introducing the role of the MAC within the risk assessment process. An expanded version of this statement is included in the FAQs on the MAC website.
16	Specify that the MAC is not appropriate for assessing manual handling of animate loads.	5.1 & 5.2.1	Web	This change has been included in the FAQ list of the on-line guidance, but not in the MAC.
17	Specify that a complete risk assessment will require consideration of the additional factors listed in the 1992 Manual Handling Operations Regulations.	5.2.1	MAC	The highlighted section in the introduction states that use of the MAC does not comprise a full risk assessment and attention is drawn to the 1992 MHOR. It also draws attention to the need to consider individual and psychosocial issues.
18	Specify the need to consult employees and safety representatives.	5.2.1	MAC	The introduction now contains a statement that refers to consulting employees and safety representatives during the assessment process.
19	Provide clear instruction on how to proceed through the flowcharts, descriptive paragraphs and score sheet of the MAC.	5.2.1	MAC	The introduction now provides step-by-step instructions on how to complete a MAC assessment.
20	Provide guidance on the possible use of individual risk bands and the total score.	5.2.2	MAC	The instructions on how to complete a MAC assessment provide guidance on these issues.
21	Mention that although the assignment of risk bands to individual risk factors may be useful in determining risk control options, they should not be considered in isolation.	5.2.2	MAC	The instructions on how to complete a MAC assessment states that the colour bands help determine which elements of the task require attention. Total scores help prioritise tasks for attention and help check the effectiveness of improvements.

22	Mention that all risk factors need to be considered together to develop and evaluate appropriate controls.	5.2.2	MAC	The instructions on how to complete a MAC assessment states that the total scores help prioritise tasks for attention and help check the effectiveness of improvements risk factors.
23	Develop interactive electronic versions of the MAC for use on personal computers, personal digital assistants (PDAs) and online.	5.3	Defer	The MAC is freely available on-line, and is supported by additional guidance which will be continually updated. If there is sufficient demand for creating additional electronic formats for the MAC, this will be considered when developing future versions of the MAC.
24	Determine if the MAC is compatible with people with defective colour vision, particularly those with reduced spectral colour discrimination.	5.3.1	MAC	The MAC has been tested with people with defective colour vision and subjective responses suggest that the colours and shades used are appropriate. In addition, alphanumeric coding has been used throughout to ensure that colour discrimination is not a prerequisite to using the MAC. In fact, the MAC has been designed to be fully functional even if printed or photocopied purely in greyscale.
25	Adapt the score sheet to allow separate scoring of lifting, carrying or team handling operations by introducing three columns.	5.3.2	MAC	The score sheet now includes three columns.
26	Shade the risk factors on the score sheet that do not apply to each operation.	5.3.2	MAC	The score sheet now includes appropriate shading.
27	Rearrange the order of the risk factors on the flowcharts to match the score sheet.	5.3.2	MAC	The order of the flow chart risk factors now matches the order on the score sheet.
28	Use a labelling system in which each risk factor corresponds to a single letter descriptor i.e. adopt one-to-one mapping.	5.3.3	Defer	This would only be feasible with a one-page flow chart format.
29	Minimise the need to flip between pages when recording scores on the score sheet.	5.3.3 & 5.3.4	Defer	Matching the order of the risk factors between the flow charts and score sheet reduces some of the flipping. A one page flow chart and score sheet would reduce flipping to a minimal level.
30	Adopt, test, and refine a single page format for the MAC flowcharts.	5.3.4	Defer	Further usability testing will be performed on this risk factor when developing future versions of the MAC.

7 REFERENCES

1 Care, B., Quarrie, C. and Monnington, S.C. (2002). *Testing and improving the usability of the Manual handling* Assessment Charts (MAC*)*. (Bootle: Health and Safety Executive). Internal report.

2 Dickinson, C.E., Riley, D.B., Boocock, M.G. and Pinder, A.D.J. (1999), *Manual handling assessment methods: A review prepared for Health Directorate D2 for issue to Regional Support Groups.* (Sheffield: Health and Safety Laboratory), HSL Internal Report EWP/99/23.

3 Jones, J.A.R., Cockcroft, A. and Richardson, B. (1999). The ability of non-ergonomists in the health care setting to make manual handling risk assessments and implement changes. *Applied Ergonomics*, **30**, 159 – 166.

4 Health and Safety Executive (1998), *Manual Handling. Manual Handling Operations Regulations 1992. Guidance on Regulations*, (Sudbury, Suffolk: HSE Books), L23, Second edition

5 Health and Safety Executive and Health and Safety Laboratory (2002), *Manual Handling Assessment Charts*. (Bootle, Health and Safety Executive), MISC 480

6 Health and Safety Executive and Health and Safety Laboratory (2003), *Manual Handling Assessment Charts*. (Sudbury, Suffolk: HSE Books), INDG383

7 Monnington, S.C., Pinder, A.D.J., and Quarrie, C. (2002). *Development of an inspection tool for manual handling risk assessment*. (Sheffield: Health and Safety Laboratory). HSL Internal Report ERG/02/20; HSL Report HSL/2002/30

8 Pinder, A.D.J. (2002), *Benchmarking of the Manual Handling assessment Charts (MAC)*. (Sheffield: Health and Safety Laboratory), HSL Internal Report ERG/02/21; HSL Report HSL/2002/31

9 Snook, S.H. and Ciriello, V.M. (1991), The design of manual handling tasks: Revised tables of maximum acceptable weights and forces. *Ergonomics*, **34**, (9), 1197-1213.

10 Tapley, S.E. (2002). *Reliability of Manual handling Assessment Charts (MAC) developed for health and safety inspectors in the UK. A field study*. (Bootle: Health and Safety Executive). Internal report.

11 Troup, J.D.G. and Rauhala, H.H. (1987). Ergonomics and training. *International Journal of Nursing Studies*, **24** (4), 325 – 330.

12 Waters, T.R., Putz-Anderson, V. and Garg, A. (1994), *Applications Manual for the Revised NIOSH Lifting Equation*. (Cincinnati, Ohio: U.S. DHHS).

8 APPENDICES

8.1 APPENDIX A – TABLES OF Z-SCORES AND P-VALUES

Table 12 z-scores and corresponding p-values for differences between briefed (n = 45) and non-briefed (n = 55) users in risk level scoring

Risk factor	*Task 1*	*Task 2*	*Task 3*	*Task 4*
Load weight / frequency	z = -1.953 p = 0.051	**z = -2.234** **p = 0.025**	z = -1.286 p = 0.199	z = -0.227 p = 0.820
Hand distance from lower back	**z = -2.470** **p = 0.014**	z = -0.371 p = 0.710	z = -0.338 p = 0.735	z = -1.754 p = 0.079
Vertical lift region	z = -0.615 p = 0.539	z = -0.295 p = 0.768		**z = -2.712** **p = 0.007**
Trunk twisting / sideways bending	z = -1.019 p = 0.308	**z = -2.230** **p = 0.026**		**z = -2.924** **p = 0.003**
Postural constraints	**z = -2.997** **p = 0.003**	z = -1.862 p = 0.063	z = -1.767 p = 0.077	z = -1.134 p = 0.257
Grip on load	**z = -3.216** **p = 0.001**	z = -0.491 p = 0.623	z = -0.067 p = 0.946	z = -0.244 p = 0.807
Floor surface	z = -0.783 p = 0.434	z = -0.589 p = 0.556	**z = -2.081** **p = 0.037**	**z = -2.264** **p = 0.024**
Other environmental factors	**z = -2.194** **p = 0.028**	z = -1.770 p = 0.077	z = -1.600 p = 0.110	z = -0.096 p = 0.924
Carry distance			z = -0.536 p = 0.592	
Asymmetrical trunk / load			z = -0.104 p = 0.917	
Obstacles en route			z = -1.367 p = 0.171	
Communication and coordination				z = -0.937 p = 0.349

Bold text indicates a significant difference ($p < 0.05$).

Table 13 z-scores and corresponding p-values for differences between briefed (n = 45) and non-briefed (n = 55) users in risk factor usability rating

Risk factor	*z-score*	*p-value*
Load weight / frequency	z = -0.587	p = 0.557
Hand distance from lower back	z =-1.625	p = 0.104
Vertical lift region	**z = -3.067**	**p = 0.002**
Trunk twisting / sideways bending	**z = -2.445**	**p = 0.014**
Postural constraints	z = -1.414	p = 0.157
Grip on load	z = -0.813	p = 0.416
Floor surface	z = -0.122	p = 0.903
Other environmental factors	**z = -3.243**	**p = 0.001**
Carry distance	z = -0.326	p = 0.744
Asymmetrical trunk / load	z = -0.720	p = 0.471
Obstacles en route	z = -0.902	p = 0.367
Communication and coordination	z = -0.159	p = 0.873

Bold text indicates a significant difference ($p < 0.05$).

Table 14 z-scores and corresponding p-values for differences between expert (n = 13) and non-expert (n = 100) users in risk level scoring

Risk factor	*Task 1*	*Task 2*	*Task 3*	*Task 4*
Load weight / frequency	z = -0.903 p = 0.336	z = -0.162 p = 0.871	z = -0.512 p = 0.609	z = -0.735 p = 0.462
Hand distance from lower back	z = -0.716 p = 0.474	**z = -4.412** **p = 0.00001**	z = -0.020 p = 0.984	z = -1.132 p = 0.258
Vertical lift region	**z = -1.959** **p = 0.050**	z = -0.878 p = 0.380		z = -0.184 p = 0.854
Trunk twisting / sideways bending	z = -0.071 p = 0.943	z = -0.234 p = 0.815		z = -0.061 p = 0.951
Postural constraints	z = -1.572 p = 0.116	z = -0.697 p = 0.486	z = -0.541 p = 0.589	z = -0.908 p = 0.364
Grip on load	**z = -2.311** **p = 0.021**	z = -0.341 p = 0.733	z = -0.915 p = 0.360	z = -1.159 p = 0.246
Floor surface	z = -0.609 p = 0.543	z = -0.625 p = 0.532	z = -1.122 p = 0.262	z = -0.250 p = 0.803
Other environmental factors	z = -0.594 p = 0.553	z = -1.465 p = 0.143	z = -0.107 p = 0.915	z = -1.823 p = 0.068
Carry distance			z = -0.335 p = 0.738	
Asymmetrical trunk / load			z = -0.472 p = 0.637	
Obstacles en route			z = -0.767 p = 0.443	
Communication and coordination				z = -0.786 p = 0.432

Bold text indicates a significant difference ($p < 0.05$).

Table 15 z-scores and corresponding p-values for differences between expert (n = 13) and non-expert (n = 100) users in risk factor usability rating

Risk Factor	*z-score*	*p-value*
Load weight / frequency	z = -1.929	p = 0.054
Hand distance from lower back	**z = -2.797**	**p = 0.005**
Vertical lift region	z = -1.411	p = 0.158
Trunk twisting / sideways bending	z = -0.786	p = 0.432
Postural constraints	z = -1.478	p = 0.139
Grip on load	z = -0.365	p = 0.715
Floor surface	z = -0.638	p = 0.523
Other environmental factors	z = -1.362	p = 0.173
Carry distance	**z = -2.749**	**p = 0.006**
Asymmetrical trunk / load	**z = -3.399**	**p = 0.001**
Obstacles en route	z = -1.244	p = 0.214
Communication and coordination	**z = -2.948**	**p = 0.003**

Bold text indicates a significant difference ($p < 0.05$).

8.2 APPENDIX B MANUAL HANDLING ASSESSMENT CHARTS (MAC) USABILITY QUESTIONNAIRE

Please answer the following questions to assist us in evaluating the effectiveness and usability of the MAC. The information that you provide will help ensure that manual handling activities can be assessed with accuracy, ease and confidence.

Please provide as much information as possible.

Many thanks for your assistance in completing this questionnaire.

BACKGROUND INFORMATION

1. **What is your role in health and safety? (e.g. health and safety manager, health and safety representative, line manager with safety responsibilities, etc.)**

2. **How much experience do you have in health and safety?** *(Please circle)*

None	Less than 1 Year	1 – 2 Years	3 – 5 Years	More than 5 Years

3. **Do you have personal responsibility for carrying out <u>manual handling</u> risk assessments?**

Yes	**No**

4. **Has the MAC been used in your workplace by an inspector? (Please circle)**

Yes	No	Unsure

How easy was it to score each risk factor? (*Please circle*)

Risk Factor	**Rating** (Very Easy … Very Difficult)	**If 4 or 5, please describe why it was difficult:**
Load Weight	**1 ---- 2 ---- 3 ---- 4 ---- 5**	
Hand Distance from the Lower Back	**1 ---- 2 ---- 3 ---- 4 ---- 5**	
Vertical Lift Region	**1 ---- 2 ---- 3 ---- 4 ---- 5**	
Trunk Twisting and Sideways Bending (*Lifting and team handling assessment only*)	**1 ---- 2 ---- 3 ---- 4 ---- 5**	
Postural Constraints	**1 ---- 2 ---- 3 ---- 4 ---- 5**	
Grip on Load	**1 ---- 2 ---- 3 ---- 4 ---- 5**	
Floor Surface	**1 ---- 2 ---- 3 ---- 4 ---- 5**	

How easy was it to score each risk factor? (*Please circle*)

Risk Factor	**Rating** (**Very Easy** … **Very Difficult**)	**If 4 or 5, please describe why it was difficult:**
Other environmental factors	**1 ---- 2 ---- 3 ---- 4 ---- 5**	
Carry distance (*Carrying assessment only*)	**1 ---- 2 ---- 3 ---- 4 ---- 5**	
Asymmetrical trunk / load (*Carrying assessment only*)	**1 ---- 2 ---- 3 ---- 4 ---- 5**	
Obstacles en route (*Carrying assessment only*)	**1 ---- 2 ---- 3 ---- 4 ---- 5**	
Communication and coordination (*Team handling assessment only*)	**1 ---- 2 ---- 3 ---- 4 ---- 5**	

5. **If available, would you use the MAC for manual handling assessment in your workplace?** (*Please circle*)

Never	Seldom	Occasionally	Often	Almost Always

6. **Are there any manual handling situations in your workplace for which you feel that the tool is not suitable for assessing the risks? If so, please describe the situation.**

7. **To what extent do you feel that use of the MAC improves your <u>confidence</u> when assessing manual handling risks?** (*Please circle*)

Does Not Improve	Slightly Improves	Improves	Greatly Improves

8. **To what extent do you feel that use of the MAC improves your <u>understanding</u> of manual handling risks?** (*Please circle*)

Does Not Improve	Slightly Improves	Improves	Greatly Improves

9. **To what extent do you feel the MAC helps to identify what <u>action to take</u> to reduce manual handling risks?** (*Please circle*)

Does Not Help	Slightly Helps	Helps	Greatly Helps

10. **Please describe any sections of the MAC that could be improved with additional written instruction. Be as specific as possible.** (*Please include page number*)

11. **What do you like or not like about the MAC?** (*You may continue writing on the reverse side of the paper, if you wish*)

Many thanks again for completing this questionnaire!