Institut de recherche
Robert-Sauvé en santé
et en sécurité du travail

Why Doesn't Training Based on Safe Handling Techniques Work?

A Critical Review of the Literature

Denys Denis
Maud Gonella
Marie Comeau
Martin Lauzier

STUDIES AND RESEARCH PROJECTS

R-1050

SUSTAINABLE PREVENTION AND WORK ENVIRONMENT

OUR RESEARCH
is working for you !

The Institut de recherche Robert-Sauvé en santé et en sécurité du travail (IRSST), established in Québec since 1980, is a scientific research organization well-known for the quality of its work and the expertise of its personnel.

Mission
To contribute, through research, to the prevention of industrial accidents and occupational diseases and to the rehabilitation of affected workers;

To disseminate knowledge and serve as a scientific reference centre and expert;

To provide the laboratory services and expertise required to support the public occupational health and safety network.

Funded by the Commission des normes, de l'équité, de la santé et de la sécurité du travail, the IRSST has a board of directors made up of an equal number of employer and worker representatives.

To find out more
Visit our Web site for complete up-to-date information about the IRSST. All our publications can be downloaded at no charge.
www.irsst.qc.ca

To obtain the latest information on the research carried out or funded by the IRSST, subscribe to our publications:

- *Prévention au travail,* the free magazine published jointly by the IRSST and the CNESST (preventionautravail.com)
- InfoIRSST, the Institute's electronic newsletter

Legal Deposit
Bibliothèque et Archives nationales du Québec
2019
ISBN : 978-2-89797-057-4
ISSN : 0820-8395

IRSST – Communications and Knowledge
Transfer Division
505 De Maisonneuve Blvd. West
Montréal, Québec
H3A 3C2
Phone: 514 288-1551
publications@irsst.qc.ca
www.irsst.qc.ca

Why Doesn't Training Based on Safe Handling Techniques Work?

A Critical Review of the Literature

Denys Denis, Maud Gonella, Marie Comeau
IRSST

Martin Lauzier
Université du Québec en Outaouais

STUDIES AND RESEARCH PROJECTS

R-1050

This study was funded by the IRSST. The conclusions and recommendations are solely those of the authors.
This publication is a translation of the French original; only the original version (R-1013) is authoritative.

PEER REVIEW
In compliance with IRSST policy, the research results published in this document have been peer-reviewed.

SUMMARY

***Background and objective*:** Training in manual material handling has been the subject of many requests in various workplaces. However, according to five meta-analyses published between 2007 and 2014, the training given to manual material handlers is of questionable value, despite being widely disseminated. A reading of these analyses does not reveal why this is so, because the training programs inventoried and evaluated were described very briefly or not at all. Having more information about handling training programs would make it easier to explain the reported lack of effectiveness and, subsequently, propose avenues for improvement. This was the objective pursued by the authors of the present study.

***Methodology*:** Seventy-seven papers covered by the five meta-analyses were analyzed using 86 variables. The training programs were first categorized according to where they took place, i.e., in the workplace, in a laboratory or in a training institution. Workplace programs were described in greater detail since they were the most numerous (51 out of 77). Categories were created based on four quality criteria supported by a theoretical framework: content adapted to context, motor engagement, contextualized practice, and workplace ergonomic transformations to complement the training. The existence of a relationship between a program’s effectiveness and the extent to which it met these criteria was verified. Lastly, a hypothesis was formulated that the meta-analyses might contain a bias related to the program selection criteria.

***Principal results*:** Training programs for manual handlers differ greatly in form, using a wide variety of measures in a broad range of contexts. The content, on the other hand, is surprisingly uniform, with a consistent emphasis on learning and adopting a safe handling technique known as “straight back, bent knees.” This standardized content is part of a training approach that focuses on the learners and their behaviours, paying little attention to the individual learner’s interaction with a changing work environment and the efforts of self-regulation this requires. In most programs, training content is predetermined and exportable from one workplace to the next, despite the differences in actual working conditions.

Of the four quality criteria accepted, only those related to transformations made concurrently to training and, to a lesser extent, training content adapted to the context, yielded improvements in terms of effectiveness. Ten percent of the studies met all the quality criteria. Despite the impressive number of studies devoted to the evaluation of handler training, the meta-analyses based their conclusions on a small number of them, assigning greater weight to those deemed to be of higher methodological quality, i.e., about one training program in ten. The results show that these higher-quality evaluation designs assessed the effectiveness of programs considered to be easy to evaluate, relatively simple, and generally of lower quality, which may have impacted the conclusions of the meta-analyses. In addition, the results regarding effectiveness, as reported by the studies' authors, paint a more optimistic picture than the conclusions reached by the authors of the meta-analyses.

***Discussion*:** The limitations of existing training programs are discussed and possible explanations are provided as to why they are reportedly so ineffective. It is important to emphasize that what should be questioned is not the relevance of offering training programs, but rather the type of training that is focused solely on learning and adopting safe handling techniques. These types of programs have been criticized, but arguments in their favour have also been provided. The techniques themselves should not be rejected, but should no longer

constitute the focus of training. Lastly, concrete recommendations are offered to improve material handling training programs.

The contradiction between the quality of the evaluation approach and the quality of the programs evaluated under this approach is also discussed. Arguments are presented about the need to develop appropriate evaluation methods to assess the effectiveness of programs considered to be more complex and hence of higher quality. Lastly, the limitations of this study and a conclusion are presented.

Highlights

- Currently, **a large majority of handling training programs consist in teaching a basic technique** known as "straight back, bent knees." Few alternatives are presented.

- **The effectiveness of this approach is questionable,** say the authors of the meta-analyses, basing their opinion on a limited number of studies whose evaluation designs were deemed to be of high quality, that is, about one in ten training programs.

- By comparison, **the studies report positive effects from more than half the training programs examined,** which contradicts the conclusions of the meta-analyses. The "straight back, bent knees" technique therefore does seem to play a certain role in accident prevention.

- **There is a risk that the meta-analyses were biased in their selection of the study samples.** Indeed, our results show that the studies using a higher-quality evaluation design expressed opinions on the effectiveness of programs deemed to be of lower quality, since they were "simpler" to evaluate.

- **This situation reflects the limitations of evaluation designs borrowed from biomedicine** (i.e., those considered to be of higher quality) when applied to more complex training given under real conditions, which is nonetheless an indication of higher quality.

- **Handling training should be re-examined** if it is to remain relevant in a comprehensive prevention program. **The content should no longer focus mainly on safe techniques, but should be rounded out with other skills specific to the handler's job context.** This context should be analyzed and understood prior to the training.

- **A skills-based approach rather than the learning of predetermined techniques is advocated.** Handlers must be able to adapt to frequent and unpredictable changes in the demands of the job.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Despite a great deal of research and intense prevention efforts both in the past and currently, manual handling remains a high-risk occupation in terms of injuries. The data show that between 2003 and 2008, more than $100 million per year was expended to compensate workers who had sustained handling accidents in Québec (Allaire and Ricard, 2007). Young workers (15 to 24 years old) are no exception, as handling is the leading cause of accidents in that age group (Ledoux and Laberge, 2006). According to recent statistics from the Commission des normes, de l’équité, de la santé et de la sécurité du travail (CNESST), there were 21,811 spinal injuries in 2010, representing 30% of all employment injuries with absence that were compensated (Provencher et al., 2011). In terms of frequency, the jobs with the highest number of spinal injuries were those of nursing assistants (2290 cases) and manual handlers (1827 cases).

The part of the body most affected is the lumbar spine (60% of spinal injuries), and overexertion was the most frequently reported source of injury, accounting for 40% of cases between 2007 and 2010. In Québec, 35,460 people have the job title of “material handler”[1] (general classification: trades, transport and equipment operators), of whom 90% are men and 10% are women (Statistics Canada, 2011). A report commissioned by the French government on labour-related needs states that, in a globalized economy, trade in goods will continue to grow, requiring an increasing number of handlers (Chardon and Estrade, 2007).

There is consensus on the importance of preventing risks associated with manual handling at their source (Rodrick and Karwowski, 2006) and which ones to target (Australian Safety Compensation Council [ASCC], 2007; Bernier et al., 2003; Mairiaux et al., 1998). The issues are not so much about what must be transformed, but about how to go about it and the possibilities for action in working environments (Denis et al., 2008; 2011). Very often, when it is impossible to provide ideal handling conditions, the workers assigned to these tasks must contend with these imperfect environments and find ways to deal with them. Generally, this reality is not unrelated to the fact that training is the most common avenue of prevention used by employers (Rivard and Lauzier, 2013). As a result, it is the focus of the prevention practices of many occupational health and safety (OHS) stakeholders and organizations.

[1] Many workers are obliged to carry out handling tasks without being in positions designated as “handler” (e.g., attendants, day labourers). Manual handling is present in numerous sectors of activity and in different forms. The term “handler” will, however, be the term used in this report to describe a person who performs manual handling activities.

2. THE ISSUE

Handling training is the subject of numerous workplace requests (Hermans et al., 2012). Despite being widely disseminated, the effectiveness of existing handling training programs has been questioned by the authors of five meta-analyses (Haslam et al., 2007; Martimo et al., 2007; Clemes et al., 2009; Verbeek et al., 2011; Hogan et al., 2014). Two of these five meta-analyses were updates of previous literature reviews, and they came to the same conclusions: no results or mixed results regarding the effectiveness of handling training. However, in spite of this observation, after consulting these reviews it was unclear why this is the case, because the training programs identified (and whose effectiveness was being evaluated) were not described at all or the descriptions lacked detail. This leads to questions about the content of these training programs, how they are taught, in which contexts and over what period of time, and whether or not they are similar. Having more information about the characteristics of handling training programs would certainly help to better explain their reported lack of effectiveness and provide avenues that could help to improve them. These clarifications would afford the opportunity to reflect on the specific characteristics of workplace training programs required for handling.

When a previous review of the literature was undertaken several years ago on interventions to prevent musculoskeletal disorders (MSDs) (Denis et al., 2008), this research team discovered that it was difficult to find an exhaustive approach combined with a proven assessment phase in the same intervention. It appeared that the more energy that was invested in diagnosis and transformation, the less energy there was for assessment. However, the opposite is also true and is most often the norm than the exception: a rigorous evaluation is commonly associated with a less complex intervention, which has a less detailed, and even sometimes inexistent, analysis of the tasks. Interventions that focus on a thorough activity analysis and that result in multiple transformations (a guarantee of their quality) use impact assessments that do not meet the criteria of standard intervention reviews. It must be kept in mind that the objective pursued by the authors of these reviews is essentially to verify the effectiveness of interventions in reducing MSDs. To reach their conclusions, these authors favour an evaluation process conducted according to a model that is as close as possible to an experimental or biomedical model. It can be assumed that the interventions selected are those for which the approach is easier to evaluate. That raises questions about the value of the findings of these reviews, whether the most appropriate approaches been considered and if it is realistic to carry out both a complex intervention and a rigorous evaluation.

On the basis of the above, the hypothesis formulated is that a similar phenomenon can be observed in reviews about the effectiveness of handling training programs. If only studies for which the assessment mechanism is experimental or biomedical (e.g., control groups, randomized trials, pre/post evaluations with longitudinal follow-up) are considered, are the training programs with the greatest OHS potential ignored? This study re-focuses the debate on the type of training provided, and not on the training as such. Only a typological analysis will make it possible to assess the handling training interventions selected in the reviews.

3. OBJECTIVES

The main purpose of this study is to understand and explain why the effectiveness of current handling training programs is under question. How did the authors of these meta-analyses arrive at their conclusions, when the dissemination of this type of training has been, and still is, so important? To answer these questions, three objectives were established:

a. This finding of ineffectiveness or partial effectiveness of training programs is based on insufficient information with respect to the characteristics of the training programs evaluated. The first objective is therefore to describe these handling training programs: how do they train people who must perform handling tasks? What are the preferred content and training systems/methods used? How is this content determined? Is it adapted to the context in which it will ultimately be used? Among other things, there is an important question concerning the diversity of types of training programs: is there one dominant paradigm or a plurality of approaches?

b. With few exceptions, the assessment mechanisms of the meta-analyses assumed that all of the training programs were basically the same. However, these training programs do not all appear to be identical and their level of quality varies. The second objective, complementary to the first, is therefore to assess the quality of training programs offered: do these training programs respect the principles recognized in the scientific literature as promoting learning? Do some of the training programs have a higher threshold of quality and effectiveness, and if so, do they achieve better results? For example, since handling is primarily a manual task in which know-how predominates, it is commonly agreed that practice (i.e., motor engagement or repetition) will lead to learning these motor skills. However, what place is given to practicing skills in training programs? How much time is allocated and according to what procedures? Do those that encourage this motor engagement achieve better results?

c. Finally, when this study began, it was hypothesized that a bias in the selection of training programs in the meta-analyses results in an inverse relationship between the robustness of the "evaluation" component and the complexity of the training interventions considered. To ensure the best evaluation, training must be simple and uncomplicated. However, if this self-selection bias in training programs proves to be justified, it necessarily has an influence on the conclusions advanced by the meta-analyses. The third and final objective is to assess whether this hypothesis is valid.

This review will result in guiding the debate on these issues in a different direction, by centring discussion on the quality of training programs provided. By providing a more accurate picture of what effects can be anticipated from a handling training program, some of the characteristics of training sessions that should be implemented to reach realistic prevention objectives in line with the expectations expressed in workplaces and by prevention experts in this field could be highlighted.

4. APPROACH

4.1 Selection of Material for Analyses

The articles selected for this review came from the five reviews of the literature (or meta-analyses) published between 2007 and 2014 on the effects of handling training. To the authors' knowledge, these five reviews are the only ones published on the subject since the beginning of the 2000s: they have therefore all been selected as basic material. Two of these reviews are from the Health and Safety Executive (HSE) in England (Haslam et al., 2007; Clemes et al., 2010), two others are from the Cochrane group (Martimo et al., 2007; Verbeek et al., 2011), and the last one is from an Irish team that used the same methodological foundation as that of the HSE (Hogan et al., 2014). The Clemes et al. (2009) review is an update of the Haslam et al. review from 2007, while the Verbeek et al. (2011) review is an update of the review by Martimo et al. (2007). Some details about these five reviews are presented in Table 4.1, including the results obtained (last column: readers can refer to Appendix 1 for more details on the conclusions formulated in the meta-analyses).

Table 4.1 Major characteristics of the literature reviews studied

Reviews	Date	No. of articles	Years covered	Objectives	Results
HSE Group					
Haslam et al.	2007	84	1980 to 2006	Handling training program effectiveness	Mixed[a]
Clemes et al.	2009	53	1980 to 2009	Handling training program effectiveness in reducing back pain and injuries	Mixed
Cochrane Group					
Martimo et al.	2007	12	1981 to 2005	Effectiveness of handling training programs and handling assistance device use in preventing and reducing back pain	Nil
Verbeek et al.	2011	18	1981 to 2010	Effectiveness of handling training programs and handling assistance device use in preventing and reducing back pain	Nil
Other					
Hogan et al.	2014	13	Before 2013	Effectiveness of handling training programs in improving knowledge and operating procedures, preventing and reducing MSDs	Nil

[a] Combination of positive and nil results or little or no "strong" evidence to conclude (evidence ranges from little to moderate, depending on the assessment of study quality).

These reviews started by using inclusion criteria (Appendix 2) to select the articles. They vary slightly from one review to the other: for example, the location where the training took place (considering workplace training programs only), the sources of published studies or the objectives arising from the training. However, after this initial selection, an evaluation of the methodological quality of the studies was conducted with respect to the evaluation approach used to assess the effectiveness of training. Thus, a lower-quality study will be assigned less weight in the conclusions formulated by the reviews and vice versa. The following section (4.2) details the steps taken by these meta-analyses to formulate their conclusions. All of the articles identified in these reviews served as the basis for this study. The study by Haslam et al. (2007), the oldest and most referenced, was chosen as the starting point: the missing references from other reviews were added over time (Figure 4.1).

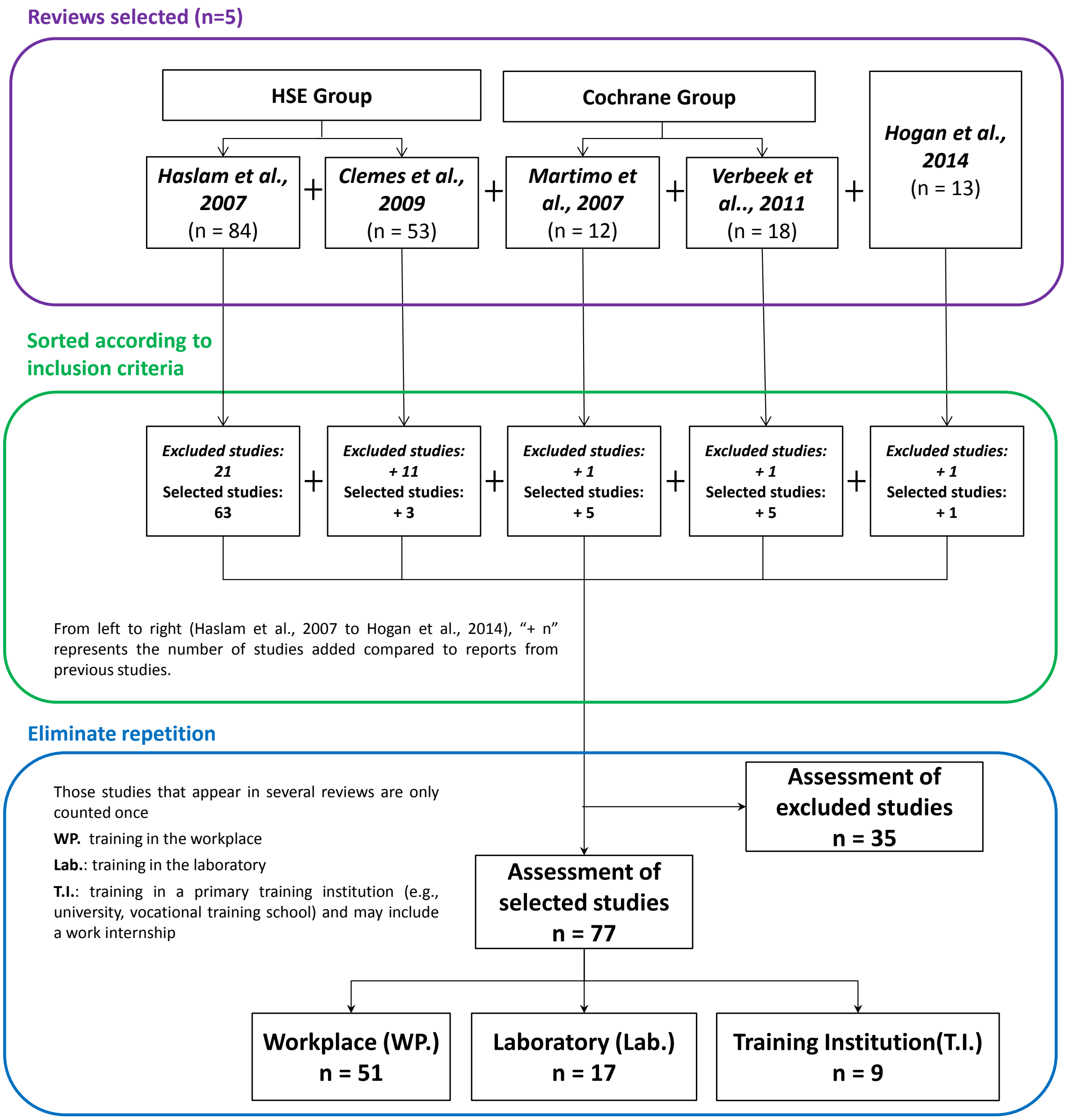

Figure 4.1 Selection process based on articles referenced in the reviews

By adding up the references provided in the five reviews, 180 documents were catalogued. The targeting steps carried out are as follows:

- removal of duplicates;
- exclusion of documents in languages other than English and French;
- exclusion of documents such as reviews or commentaries from experts;
- documents that cannot be traced (references not mentioned or impossible to find).

After making the selection, 77 studies in which training programs had been provided in three separate locations were accepted:

- **In the workplace (WP.)** (n=51): these training programs represented two thirds of our sample. They took place in the workplace within a variety of organizations (e.g., hospitals, storage and delivery companies). This was the type of handling training that is most representative of the practice of prevention experts;
- **In the laboratory (Lab.)** (n=17): this environment, most closely associated with experimental research, is "controlled" (e.g., biomechanical, physiological laboratories), which makes it possible to perform complex measurements (e.g., EMG, kinematics) that are difficult or impossible to carry out in a real work environment. Although it facilitates the control of independent variables and data collection, the laboratory environment is less representative of workplaces, because it is usually simplified (e.g., the types of load, force platforms that do not enable mobility of the feet, simple lifting tasks);
- **In training institutions (T.I.)** (n=9): this refers to the initial (qualifying) training received in educational institutions (e.g., vocational schools) by future employees who will have to perform handling tasks (e.g., nurses, attendants). Manual handling is generally a module in their curriculum and the training may include a work internship.

The articles selected by the HSE reviews include studies conducted in laboratories, in the workplace, or in training institutions, while those of the Cochrane group were conducted only within the workplace. The most recent review (Hogan et al., 2014) enabled us to add only one new article compared to the other four. Finally, it should be noted that for the Clemens et al. (2009) review, 11 of the 53 articles used were not clearly referenced, which meant they were impossible to find.

4.2 Methodology Used in the Meta-analyses

In order to better understand the methodology used by each review to formulate their conclusions, the reference documents they cited were consulted: some of them, however, provided very little detail about certain steps in the analysis. Despite the high degree of rigour that characterized these meta-analyses in their evaluation process, it was not easy to understand how they were rated (see Appendix 6 for examples of contradictory classifications) or how they were subsequently used to draw conclusions. The lines that framed their approach are presented as they appear in the documents consulted.

Evaluation of the methodological quality of studies: All the reviews evaluated the methodological quality of the selected studies. To do so, the Cochrane group (Verbeek et al., 2011; Martimo et al., 2007) used the guidelines recommended by the Cochrane Back Review Group for randomized controlled trials (RCT) (Clarke and Oxman, 2002; Higgins and Green, 2008). Verbeek et al. (2011) used the Furlan (2009) version, which is an update of the van Tulder (2003) version, and which was also used by Martimo et al. (2007). Depending on the year of publication, the guidelines include a list of 11 or 12 questions to determine the risk of bias of the RCTs (e.g., is the randomization method appropriate?). With respect to the specific case of cohort studies, the Cochrane group used the methodological index for non-randomized studies (MINORS), developed by Slim et al. (2003), which includes 8 to 12 items, depending on the types of studies analyzed. These items include the use of longitudinal follow-up adapted to the objective and the realization of appropriate statistical analyses. However cohort studies were used only to compare RCTs, with no impact on the Cochrane group's overall conclusions.

The HSE group (Haslam et al., 2007; Clemes et al., 2009) and Hogan et al. (2014) used the Downs and Black (1998) checklist to assess the methodological quality of the studies. This checklist was originally developed to evaluate studies in the health sector. It includes 27 questions about elements included in the study (e.g., clarity of the hypothesis or the objective), its external validity (e.g., the representativeness of the sample), internal validity (e.g., the sampling strategy) and statistical power. Hogan et al. (2014) used a modified version of this checklist that adjusts the rating for one question (#27: statistical power), while the HSE group reviews added two other questions to that checklist, namely, the presence of a control group and the realization of a longitudinal follow-up (follow-up period).

Interpretation of results and conclusions: In order to formulate their conclusions, the reviews used the methodology recommended by the Cochrane Back Review Group. Hogan et al. (2014) reported using the same methodology as the HSE group. They used the van Tulder (2003)[2] version, which based the strength of the scientific evidence on the methodological quality and the results of the studies. The latter were weighted using statistical analyses such as calculation of the odds ratio, the mean deviation, and/or the standard deviation. For example, the strength of evidence would be high if two or more studies of high methodological quality had statistically significant results in the same direction for 75% of the trials reported (Martimo et al., 2007, pp. 5-6):

> *"A qualitative analysis was completed using a rating system, based on levels of evidence, to summarize the strength of scientific evidence of the effects of the intervention. The rating system was based on both the quality and the outcome of the studies (van Tulder, 2003):*
>
> *I. Strong evidence—consistent evidence in two or more high quality RCTs;*
> *II. Moderate evidence—consistent findings in multiple low quality RCTs and/or one high quality RCT;*
> *III. Limited—one low quality RCT;*
> *IV. Conflicting evidence—inconsistent findings in multiple RCTs;*
> *V. No evidence—no RCTs.*

[2] Haslam et al. (2007) do not report how they arrived at their conclusions. Because they are part of the HSE group, it was inferred that they used the same approach as Clemes et al. (2009).

> *The outcome of the studies was considered 'consistent' if at least 75% of the trials reported statistically significant results in the same direction."*

Martimo et al. (2007) also used the van Tulder (2003) version. However, Verbeek et al. (2011) used the Furlan (2009) version, which seems to add steps to the definition of scientific evidence with respect to van Tulder (2003). The strength of the evidence is based on five criteria:

(1) *Limitations*: risk of bias (assessment of methodological quality);
(2) *Inconsistency*: absence of similarities/consistency between the results and estimates of effects measured across the studies (consistency reached when direction, size effects and statistical results lead to the same and statistically significant conclusions);
(3) *Indirectness*: difficulty in generalizing the results (population, intervention and outcomes that are not comparable to those considered by the inclusion criteria in the review of the literature)
(4) *Imprecision*: too great of a margin of error/confidence interval measured (e.g., median deviation, standard deviation, etc.) for each result (e.g., number of participants, events)
(5) *Publication bias*: probability of the selection of trials and results presented (e.g., results expected and presented in the protocol, but not presented in the results).

For each criteria that is not met, the strength of the evidence decreases incrementally (Furlan et al., 2009):

> *"High quality evidence = at least 75% of the RCTs with no limitations of study design have consistent findings, direct and precise data and no known or suspected publication biases.*
> *Moderate quality evidence = 1 of the domains is not met.*
> *Low quality evidence = 2 of the domains are not met.*
> *Very low quality evidence = 3 of the domains are not met."*

As Martimo et al. (2007) report, if the evaluation of the methodological quality is combined with the results measured by the studies to formulate conclusions, the higher the methodological quality of a study, the more weight it has in the formulation of conclusions. <u>This is why the methodologies used by the studies hold a dominant place in the analysis conducted by the reviews</u>. The only elements related to the intervention appear in the initial inclusion criteria of the articles (e.g., a training program conducted in the workplace or laboratory, physical conditioning, training in manual handling). After that, the type of intervention holds no weight in the conclusions reached.

As a result of these analyses, the reviews conclude that **handling training programs have no impact on preventing back pain and injuries, reducing work-related musculoskeletal disorders and on changing behaviours**. However, some saw beneficial effects in the form of reducing the risk of handling-related accidents through (short-term) physical conditioning or "ergonomic" training programs/interventions.

4.3 Analyses Undertaken

4.3.1 General Description of Study Content

In order to describe the content of the 77 articles selected, five levels of analysis were identified and used following a reading of a random sample of the articles in order to summarize the information (Table 4.2). Furthermore, since the research team included handling training specialists, some variables specific to each of these levels were selected, because they are characteristic of this type of training and make it possible to describe it. As coding progressed, where necessary, new variables were added according to the material available: the ergonomist then performed iterations to ensure that everything was covered.

Table 4.2 Examples of variables according to analysis level

Description of analysis levels	No. of variables	Examples of variables
Study: this first level, the most macroscopic, includes the entire research project, from beginning to end: the timespan extends from the intervention to its assessment.	3	Objectives pursued Number of objectives pursued Overall duration of the study
Intervention: corresponds to all of the actions performed during the process, including training and transformations, where applicable.	27	Sector of activity Workplace size Size and type of intervention Duration of the intervention Reported phases Participants' characteristics
Training: principle activity undertaken by the participants: physical training, learning safe techniques, etc.	32	Profile of trainers Size of learning group Location and format of sessions Duration and time spread Types of loads used Educational methods and tools Content and knowledge transmitted
Transformations: generally as an adjunct to training, modifications to the initial work situation: handling aids, layout, PPE, etc.	1	Presence or absence of transformations
Evaluation: the entire process used to assess the impact of the intervention, its effects	23	Evaluation design Collection and latency period Tools and indicators used Changes measured

For each of these levels, variables (between 1 and 32, for a total of 86 variables) were selected (see Appendix 3 for the details about the definitions of each variable. This appendix also contains the variables used in the meta-analyses: n=16, for a total of 110 variables). The information was collated by an ergonomist using NVivo® software. Most of the variables had to have been clearly provided by the authors, while others were inferred from a variety of information presented by the authors (such as the training paradigm used). The analyst only made a decision if the information was clear and consistent.

Processing by occurrence and percentage was then carried out. The information available in the studies will be reported first and the major characteristics of the training approaches will be described, depending on the three locations in which they took place. Afterward, workplace training will be examined separately and described in detail.

4.3.2 Training Categorization Trials

The meta-analyses ranked studies according to the quality of their evaluation approach: they made no judgment about the quality of the training itself, but concentrated on the quality of the design used to assess its effectiveness. This study has instead the objective of ranking (or grouping) training according to the quality of the approach,[3] in order to determine whether these groups differ in terms of effectiveness (4.3.2.2): is training that is deemed to be of high quality more effective? To rank the programs, training quality criteria were defined on the basis of theoretical positioning and the scientific literature associated with it (4.3.2.1).

These criteria also serve to test the hypothesis that there is a bias in the selection of training programs in the meta-analyses (4.3.2.3). Recall, for the record, the idea that a rigorous evaluation would only be possible for less complex training programs, thus programs of lower quality. However, in formulating their conclusions, the authors of the meta-analyses gave significantly more weight to studies with a high-quality evaluation design: are they therefore commenting on lower quality training programs, thus concealing the actual potential of handling training programs to reach their objective of preventing injuries? That is what this study seeks to verify.

4.3.2.1 The Four Criteria Selected[4] to Evaluate the Quality of a Training Program

Training approaches arise from very diversified currents of thought and are based on learning theories that are just as varied. Researchers interested in training issues generally confine themselves to one and use theoretical foundations to justify their position. The authors of this report follow a training approach based on skills development (Le Boterf, 2003), as these skills help workers regulate the variable work situations they face by adapting their work activity to the situation (Guérin et al., 2007; St-Vincent et al., 2014).

In this social constructivist-inspired approach (Jonnaert, 2009), it is necessary, but not enough for learners to simply acquire knowledge and expertise at the end of a training program. They must also, and above all, be able to mobilize what they know to solve problems related to their work situation: what should I do in this particular situation to solve the problem and achieve my

[3] The reader should make a clear distinction between the quality of the training (which is what we are trying to highlight here) and the quality of the design or methodology used to demonstrate the effectiveness of that training (the quality of the evaluation). Meta-analyses focus only on the quality of the evaluation, without considering the quality of the training they evaluate.

[4] The list of criteria used is not exhaustive: these four criteria have been identified both because they are representative of the authors' educational philosophy and because it was possible to find data concerning them in the studies considered in this review. To illustrate, the quality of feedback during learning activities could have been used as a relevant criterion, but too few studies provide sufficient detail about this dimension, which is considered essential to learning in general and for motor learning in particular. (Schmidt, 1988).

goals? Individuals who are able to select and apply knowledge relevant to the problems they are facing in their environment are qualified as being competent. To be useful to them, the knowledge they put into action must therefore be linked to the contexts in which they are used—to their specificities—and be enough to adapt to most problematic configurations found in work situations.

First criterion selected: training content must be adapted to the context in which it will be used. Prescribed and general knowledge and/or standardized and predefined working methods may be taught, but they will not be sufficient to adapt to all the possible configurations that work situations may take: they must be completed by contextualized practical knowledge. To identify the latter, preliminary analyses of work situations concerned by the training program appear to be essential.

Once the training content has been established, the way it will be taught to learners must be determined: this is what is commonly referred to as the educational approach. To judge the quality of choices related to the approach in the logic of competency-based training, the concept of self-regulation is central. Work situations are not static; they transform with the disruptions and uncertainties that occur in the workplace. Workers must therefore constantly adjust to the dynamic/changeable character of situations they are experiencing: this is what has been defined as a person's competence.

It is this competence that makes self-regulation possible. Keep in mind that handling is above all a manual activity: the motor skills that will be mobilized must be mastered (e.g., knowing how to pick up a load, knowing where to place one's feet to maintain good balance). Workers do not pick up all loads in the same way: they adapt themselves according to the weight, fragility, volume or the presence of handles on the object, for example. It is here that the individual's competence takes on its full meaning: people will adapt their way of doing things to lift a load according to their understanding of its characteristics.[5]

Two major training challenges emerge here, to which the approach must be adapted. First, the person must learn how to grasp a load: where should the hands be placed on it? Which hand will support the weight? Which will guide the movement? What level of pressure is required to control it effectively? Repeating the movements enables progressive mastery of the task. The example of grasping is easy for the reader to understand, but is certainly not the most complex skill to be learned in handling: maintaining one's balance throughout all the phases of handling is a more difficult motor learning process.

Second criterion selected: training must provide time for workers to apply the knowledge, and especially the skills, that they must master: there must be motor engagement by the trainee. This is particularly true when learning how to perform a movement in which repetition is considered as fundamental to mastering it. (Schmidt, 1988).

Next, the person must be exposed to loads with diverse characteristics, which are representative of what must be handled on a daily basis, in order to develop the know-how to

[5] With regard to the first criterion, a training program that only presents one way of grasping a load would not enable a learner to develop skills in that area; it will only make the person capable of executing the same standardized grasping technique, regardless of the variability of the loads.

choose the type of grasp best suited to the particular object: a 25 kg bag of flour is not picked up in the same way as a 10 kg case of water bottles or a 50 kg table. The more variable the loads are in the workplace, the higher the self-regulation/adjustment requirements will be.

> ***Third criterion selected:*** the practice necessary to learn the movements required for handling must take place in an environment representative of the situations in which the workers will find themselves once their training is completed. To return to the example of types of grasps: why practice how to grasp boxes if they will not be the type of object that will be handled after the return to work? This criterion is related to the concept of "situated learning" from the social constructivist theory of learning. Context-based learning gives meaning to learners' strategies for building knowledge and developing skills: the link between what people want them to learn and its usefulness in a real context will be much clearer to them. We know that the usefulness perceived by learners of the knowledge being taught is a factor that explains how they transfer their learning into context (Lauzier and Denis, 2016). In such an approach, meaningful learning and the transformation of new information into viable and transferable knowledge are emphasized.

In this training logic, it is quite clear that it is not solely individuals and their ways of doing things that are at the centre of learning, but instead, their interactions with their environment and how they adapt their behaviour to the context. It is the problems posed by this environment that people in training must learn to solve. It is logical to assume that if one were able to reduce the level of difficulty associated with a given work context, training would be facilitated. In the short term, the trainers' role is to control the level of difficulty of the situations chosen in training and to adapt it to the trainees' skills. In the medium and longer term, transforming work situations sustainably to keep them from being too restrictive/difficult can therefore prove to be very beneficial as a complement to training: lower levels of skills will be required to regulate these situations.

> ***Fourth criterion suggested***: in addition to training activities, transformations to work situations should be encouraged to reduce the constraints. These transformations should not be wholly focused on the individual (e.g., on protective equipment or handling devices), but should above all target the aspects from which restrictions arise: well-designed work layout (e.g., appropriate heights for picking up objects and setting them down), load weights that respect standards, uncluttered workspaces, etc. Adapting the work context will enable the use of more efficient handling techniques in terms of OHS and productivity.

This last criterion stands out from the others in that it is generally not an integral component of training programs; it is instead the subject of parallel activities. It can be seen as a highly desirable addition to a training approach, which we will refer to as a "training intervention." Unlike this criterion, the three others are directly associated with training. Moreover, for a training program to be seen as being of high quality (according to the theoretical framework adopted in the context of this study), it should include each of these criteria because they are interdependent and interrelated. The presence of any of these criteria could potentially increase the quality of a training program, but it is above all their combination that makes a training program comprehensible and relevant.

Improving material handlers’ skills therefore involves the following cascade of actions: ***a.*** first, choosing training content based on the realities of the context in which it will be used; ***b.*** second, putting the elements of this training content into concrete practice by the participants, especially with respect to motor skills; ***c.*** third, this motor engagement should take place in a context that is as close as possible to the situation in which the learner will work at the end of the training session. This chain is sequential and any missing link will diminish the quality of training, especially if the first criterion is not met: why practice something if the worker is unlikely to use in his or her work context? Keep in mind that these are occupational training programs.

In conclusion, the questions related to setting objectives that should be reached at the end of the training program are essential. Handling training generally has the goal of preventing MSDs and associated symptoms: that is the definitive or final objective. Intermediary objectives can, of course, be determined: these objectives represent steps or levels to reaching the ultimate goal of preventing MSDs. Kirkpatrick (1994) developed a model to evaluate training activities in which he determines the cascade of objectives to be reached to arrive at the expected conclusion.

The first level concerns participants’ reactions: did they appreciate the training program? According to Kirkpatrick, this is a prerequisite for reaching the second level, that of learning: did they acquire knowledge? The second level leads to a third and last step before the final objective is reached. It has to do with behaviour: did the participants use what they learned when they returned to work? This step is commonly referred to as the “transfer of knowledge or learning.” To sum up, if the participants did not find the training interesting, if they were not motivated and/or they did not see the information taught as being useful to them (all possible reactions to training), according to Kirkpatrick, it is unlikely that they will be able to complete the second step and learn something meaningful: the same is true for the other steps. Here we see the importance of the first criterion: content that is too far removed from the daily reality of participants can have a negative impact on their motivation, and this reaction can be detrimental to learning.

Each objective of this sequence presents educational challenges that should not be underestimated. However, and without making it a criterion of quality, it is felt that the higher the specific purpose of a training program in this hierarchy of objectives outlined above, the more likely it is to prevent MSDs related to handling loads: wanting the participants to adopt the desired behaviour upon their return to work is a more ambitious objective than acquiring knowledge.

4.3.2.2 Group Types for Workplace Training Programs

Training programs were then classified in consideration of these four criteria. This classification only concerns studies that took place within organizations, because they are more directly related to vocational training, which is the subject of most requests from employers. Of the 51 studies conducted within organizations, three were withdrawn from the analyses because they were exclusively focused on the development of physical capacities, resulting in a total of 48 studies (> 60% of training programs identified).

Four types of groups were established, one for each criterion retained to judge the quality of a training program. An initial group made it possible, through an analysis of how well matched the training content was with workplace realities, to compare a group that had received training with content adapted to their work situation to another group for whom the content was not adapted.

The three other quality criteria for mainly manual-dominant training were treated identically to the first: whether or not participants are offered a practical component to encourage their motor engagement; whether or not participants are able to practice in a context that is representative of their jobs; and, whether or not the work situation is transformed as an adjunct to training.

These four classification attempts were cross-referenced with the variables used to assess the effectiveness of the training programs. Effectiveness was determined from the conclusions reported by the authors of the studies, regardless of the quality of the design used to formulate them. The intention was to verify whether meeting or not meeting these criteria has an impact on the effectiveness of a training program.

4.3.2.3 Verification of the Meta-analysis Selection Bias Hypothesis

To test the hypothesis that there is an inverse relationship between the robustness of the methodology used to assess the effectiveness of training programs and their quality, the four criteria regarding quality were used again. First, the studies identified in the meta-analyses with the highest rating in terms of the quality for their evaluation designs were selected. On the basis of the four criteria, those studies were compared to two other groups of studies: those rated as having a good to average quality of methodology (grouped together under the "other" category) and those with the lowest scores. A "methodology quality gradient," was obtained, ranging from the highest to the lowest rated studies in the meta-analyses.

To confirm the hypothesis, an inverse relationship must be found in terms of the quality of the training approach: the higher the quality of the evaluation design, the lower the quality of this group's training program, according to our four criteria. If this relationship is confirmed, one can acknowledge that the meta-analyses have biases in the selection of training programs. Of course, the intention is not to blame them for only selecting high quality designs, but rather to establish that, by so doing, the training programs for which one would have expected the best effects are not considered at their fair value when comes to conclusions about their effectiveness, because of an inability to properly assess them.

5. RESULTS

This section is divided into three parts, each corresponding to the objectives of this study. The first part (5.1), reports on the key characteristics of handling training programs, according to the three locations where they took place, as well as any information that was available in the studies analyzed. Afterward, a much more detailed portrait of training processes is provided, but only for workplace training (5.2). This choice is justified, on the one hand, by the fact that those training programs represent two thirds of all the training programs identified (51/77) and, on the other, because they constitute the type of training most representative of the practice of prevention experts. The last part is devoted to verifying the central hypothesis of this study: whether there is a selection bias in the meta-analyses that assess the effectiveness of handling training programs (5.3).

5.1 Description of Handling Training Program Characteristics

In this first part, the type of information to which we had access in the studies chosen will be discussed (5.1.1). Afterward, a description of the main characteristics of these training programs will be outlined (5.1.2). The results presented here concern all of the 77 studies selected and are broken down according to the three locations in which the training programs identified took place. To conclude this section, a detailed account of the 51 workplace-training programs will be presented (5.1.3).

5.1.1 Information Available in the Studies Identified

In general, the published studies provide most of the information needed for a fairly good representation of the training program, especially in terms of how it was evaluated (Table 5.1). In fact, all the studies describe in detail the process followed to evaluate the training program. The same cannot be said for certain characteristics of the training process and the actors concerned.

In fact, few studies (< ⅓) report information about the size of the organization in which the training took place (for studies in the workplace: first column, Table 5.1), the trainers' profiles and their level of experience, and the size of groups of trainees. Many characteristics of participants in training programs were not reported in the studies (e.g., age or injury history). Similarly, little is known about the educational tools used. On the other hand, information related to the sectors of activity in which the handlers work, the objectives of the training program, the profile and occupation of the participants and even the subjects covered are usually specified. For three quarters (73%) of the training programs, the duration is known, but the information about the duration of the practical and theoretical portions of training, and how long they last (the number of days over which training sessions take place) is more fragmentary. When the training is offered in the workplace, the description of the tasks for which it was designed is present in less than one third (31%) of the studies. And even when information about the tasks is available, it is minimal and general, with the result that the characteristics of the handling tasks for which training is suggested is little known or unknown.

Table 5.1 Information found in the studies

Variable considered	WP. (n = 51)	Lab. (n = 17)	T.I. (n = 9)	Total (n = 77)
Duration of study	43 (84%)	15 (88%)	4 (44%)	62 (81%)
Duration of intervention	34 (67%)	14 (82%)	8 (89%)	56 (73%)
Objectives targeted	51 (100%)	17 (100%)	9 (100%)	77 (100%)
Sector of activity	51 (100%)	16 (94%)	9 (100%)	76 (99%)
Workplace characteristics:	51 (100%)	N/A	N/A	N/A
Workplace size	16 (31%)	N/A	N/A	N/A
Number of workplaces	49 (96%)	N/A	N/A	N/A
Tasks performed in the workplace	16 (31%)		N/A	N/A
Participants' characteristics:	51 (100%)	17 (100%)	9 (100%)	77 (100%)
Gender	31 (61%)	17 (100%)	6 (67%)	54 (70%)
% Women	25 (49%)	14 (82%)	6 (67%)	45 (58%)
Age	23 (45%)	10 (59%)	3 (33%)	36 (47%)
Seniority	10 (20%)	N/A	N/A	NA
Medical history	27 (53%)	13 (76%)	4 (44%)	44 (57%)
Profile	51 (100%)	14 (82%)	9 (100%)	74 (96%)
Occupation	50 (98%)	13 (76%)	9 (100%)	72 (94%)
Handling experience	32 (63%)	11 (65%)	5 (56%)	48 (62%)
Trainers' characteristics:	28 (55%)	4 (24%)	2 (22%)	34 (44%)
Profile	22 (43%)	2 (12%)	0 (0%)	24 (31%)
Origin	14 (27%)	N/A	N/A	N/A
Experience	6 (12%)	0 (0%)	0 (0%)	6 (8%)
Number	24 (47%)	3 (18%)	0 (0%)	27 (35%)
General training characteristics:	49 (96%)	17 (100%)	9 (100%)	75 (97%)
Size of groups	12 (24%)	4 (24%)	2 (22%)	18 (23%)
Location	42 (82%)	17 (100%)	6 (67%)	65 (84%)
Format of sessions	40 (78%)	17 (100%)	5 (56%)	62 (81%)
Total duration	31 (61%)	15 (88%)	6 (67%)	56 (73%)
Duration of theoretical portion	15 (29%)	12 (71%)	3 (33%)	30 (39%)
Duration of practical portion	20 (39%)	12 (71%)	3 (33%)	35 (45%)
Time range	28 (55%)	15 (88%)	8 (89%)	51 (66%)
Type of load used	43 (84%)	16 (94%)	7 (78%)	66 (86%)
Educational methods:	45 (88%)	17 (100%)	7 (78%)	69 (90%)
Sequence	44 (86%)	17 (100%)	7 (78%)	68 (88%)
Educational tools	25 (49%)	9 (53%)	6 (67%)	40 (52%)
Subjects covered	48 (94%)	17 (100%)	6 (67%)	71 (92%)
Evaluation characteristics:	51 (100%)	17 (100%)	9 (100%)	77 (100%)
Evaluation model	51 (100%)	17 (100%)	9 (100%)	77 (100%)
Control group	51 (100%)	17 (100%)	9 (100%)	77 (100%)
Collection time (pre-/post-/follow-up)	51 (100%)	17 (100%)	9 (100%)	77 (100%)
Evaluation tools used	51 (100%)	17 (100%)	9 (100%)	77 (100%)
Indicators used to evaluate training	51 (100%)	17 (100%)	9 (100%)	77 (100%)

< ⅓ gave information
< ½ gave information
< ⅔ gave information

5.1.2 Some of the Key Characteristics of Training Programs

The analysis of the general characteristics of the studies, according to the three main locations where the training programs took place, reveals some differences (Table 5.2). Despite the fact that all the studies concerned handling training, their objectives were not the same. While one of the objectives in most of the workplace training programs (65%) was to reduce accidents, MSDs and/or injuries, slightly over one in two laboratory studies (53%), focused on the improvement of physical capacities. In training institutions, although the objectives were more diversified, these training programs stood out from the others by concentrating on changing behaviours and enhancing knowledge.

Is there a link between improving physical capacities and preventing MSDs?

Four sources, including two literature reviews, were consulted to determine the effects of physical conditioning and improved cardiovascular or muscle capacity (flexibility, endurance, strength, etc.) in preventing MSDs, accidents or occupational diseases. Blue (1996) notes that a worker with a physical condition that is up to the demands of the job is less likely to be injured. The author was, however, unable to state whether physical condition is a preventive factor. According to two literature reviews (Tveito et al., 2004; Bigos et al., 2009), there is strong to limited evidence of the effectiveness of physical training in reducing back pain. However, Claudon et al. (2016) state that few studies show long-term effects of physical conditioning in reducing back pain. No link could be established between conditioning or improving muscle and/or cardiovascular capacities and the prevention of accidents or reduction of MSDs.

Among the studies analyzed, the health sector is by far the one in which handling training is most widely offered: all courses held in training institutions were for that sector, as was 61% of workplace training. Laboratory training was the most generalized, with 59% of studies not targeting any particular sector. As noted in Table 5.2, participants in laboratory training are among the most atypical, with most being students (young men, of whom 71% have no health problems: see Appendix 4) who did not intend to work in jobs requiring handling, unlike those who followed programs in training institutions (future nurses or attendants).

When training takes place in the workplace, it is more frequently accompanied by complementary actions such as physical conditioning or transformations to the work situation. In the laboratory, more than one in two training programs focuses on physical conditioning (53%), which is in line with the above-mentioned objective of increasing learners' physical capacities. Finally, the durations of these programs are longer for workplace and institutional training, while they run for less than six months for laboratory training. Forty-two percent of training programs had a duration that was equal to (16%) or less than (26%) one day.

Table 5.2 Key characteristics of studies

General Characteristic	WP. (n = 51)	Lab. (n = 17)	T.I. (n = 9)	Total (n = 77)
Objective – Evaluate the effectiveness of the training on[1]				
Reducing accidents/ MSDs, pain	33 (65%)	1 (6%)	4 (44%)	38 (49%)
Changing behaviours	14 (27%)	2 (12%)	5 (56%)	21 (27%)
Improving physical capacities	6 (12%)	9 (53%)	-	15 (19%)
Improving knowledge	8 (16%)	1 (6%)	5 (56%)	14 (18%)
Reducing health risk factors	6 (12%)	4 (24%)	-	10 (13%)
Other[2]	7 (14%)	2 (12%)	3 (33%)	12 (16%)
Sector of activities[1]				
Health/hospital	31 (61%)	1 (6%)	9 (100%)	41 (53%)
Other sectors (e.g., construction)	18 (35%)	5 (29%)	-	23 (30%)
No specific sector[3]	-	11 (65%)	-	11 (14%)
Multiple	2 (4%)	-	-	2 (3%)
Profile of participants[1]				
Workers	50 (98%)	5 (29%)	-	55 (71%)
Students	2 (4%)	11 (65%)	9 (100%)	22 (29%)
Managers	6 (12%)	-	-	6 (8%)
Supervisors	4 (8%)	-	-	4 (5%)
Type of intervention				
Physical conditioning only	3 (6%)	9 (53%)	-	12 (16%)
Training only:	27 (53%)	7 (41%)	9 (100%)	43 (56%)
+ physical conditioning	5 (10%)	1 (6%)	-	6 (8%)
+ change(s)	15 (29%)	-	-	15 (19%)
+ physical conditioning and change(s)	1 (2%)	-	-	1 (1%)
Duration of intervention				
≤ 1 month	8 (18%)	10 (59%)	1 (11%)	19 (25%)
1 to 6 months	14 (27%)	4 (24%)	-	18 (23%)
6 to 12 months	6 (12%)	-	1 (11%)	7 (9%)
12 to 24 months	4 (8%)	-	2 (22%)	6 (8%)
> 24 months	2 (4%)	-	5 (56%)	7 (9%)
Temporal distribution of training sessions[4, 5]				
1 day	10 (20%)	5 (29%)	2 (22%)	17 (22%)
2 to 30 days	8 (16%)	6 (35%)	-	14 (18%)
31 to 365 days	10 (20%)	4 (24%)	2 (22%)	16 (21%)
1 to 2 years	2 (4%)	-	2 (22%)	4 (5%)
More than two years	-	-	4 (44%)	4 (5%)
Total duration of the training program[5]				
< 1 day	13 (25%)	5 (29%)	2 (22%)	20 (26%)
1 day	8 (16%)	3 (18%)	1 (11%)	12 (16%)
≥ 2 days	10 (20%)	4 (24%)	2 (22%)	16 (21%)

1. Sub-categories that are not mutually exclusive
2. Other: the studies' objectives could be to evaluate the content of a training program, its impact on participation in leisure activities, its impact on changes in work activities, etc.
3. Does not target any specific activity sector (e.g., students in unspecified fields of study)
4. Some studies included more than one training program with different characteristics
The temporal distribution represents the range in time over which the training sessions take place. The duration of the training program is the actual length of time it took. For example, training with a total duration of 1 day may be broken up into more than one session held over a 30-day period.

5.1.2.1 Evaluation of the Methodological Quality of Studies and the Effectiveness of Training

In Table 5.3, two types of results are reported: the evaluation of the methodological quality carried out in the meta-analyses and the effectiveness of training as reported in the studies. The majority of training programs identified by the reviews were below the "high" quality threshold in terms of the evaluation methodology used, even in a controlled environment such as the laboratory, where all the studies were deemed to be of "other" quality, i.e., good (18%), average (35%) or low (47%). In fact, one in ten studies (10%) is considered to be of a high methodological quality, with those conducted in organizations being the most highly rated in terms of the quality of their evaluation process (14%). Note that in the studies deemed to be of "other" quality, some were still rated as being good or average: they were therefore taken into account in the reviews to evaluate the effectiveness of training, but they will have a lower weight. This is the case for more than half of the laboratory studies (good to average: 53%) that were used in some reviews to justify their conclusions. In fact, the laboratory studies that were "recovered" to formulate the authors' conclusions regarding the meta-analyses are essentially those that used a control group.

Table 5.3 Ratings of the methodological quality of the evaluation and results reported in terms of training program effectiveness

Evaluation component	WP. (n = 51)	Lab. (n = 17)	T.I. (n = 9)	Total (n = 77)
Classification of the methodological quality of articles by the reviews[1]				
High quality	7 (14%)	-	1 (11%)	8 (10%)
Other quality[2]	30 (59%)	17 (100%)	2 (22%)	49 (64%)
Good quality	8 (16%)	3 (18%)	-	11 (14%)
Average quality	6 (12%)	6 (35%)	1 (11%)	13 (17%)
Poor quality	12 (24%)	8 (47%)	1 (11%)	21 (27%)
Contradictory classification	4 (8%)	-	-	4 (5%)
Unassigned	12 (24%)	-	6 (67%)	18 (23%)
Contradictory classification	2 (4%)	-	-	2 (3%)
Results:				
Authors' assessment of the effectiveness of the activity				
Positive	25 (49%)	13 (76%)	2 (22%)	40 (52%)
Mixed	5 (10%)	1 (6%)	3 (33%)	9 (12%)
Nil	21 (41%)	3 (18%)	4 (44%)	28 (36%)

1. The studies were rated according to the most recent classification of the review and using the same methodology (e.g., Haslam et al, 2007, Clemes et al, 2011 and Hogan et al, 2014 used the same checklists to evaluate the methodological quality. As Hogan et al, 2014 is the most recent, its classification was used if there was a contradiction among the three reviews.)

The Cochrane group cohort studies can be found in this category, because they were not used by the reviews to formulate their conclusions.

The authors of the laboratory studies reported having obtained the most positive effects following training, with a very impressive rate (76%), with 53% having improved the physical capacity of trainees (see Appendix 4). In fact, all the training programs that focused on the development of physical capacities reached their objectives. Workplace training had a positive impact in half the cases, while in training institutions, 22% of the programs had a positive impact. In general, based on the effects reported by the authors of the studies, more than half of the training programs (52%) had positive results (to which can be added 12% of mixed effects), in contrast to the conclusions drawn by the meta-analyses, which took into account the robustness of the evaluation mechanisms for their judgements. Thus, the impressive results obtained in the laboratory would be put into perspective in the reviews, given the poor quality of their evaluation methodology[6].

5.1.3 Characteristics of Workplace Training

Because they represent two thirds of the training programs identified and are the training method used the most by stakeholders, the authors wanted to see what the specific features of workplace training were. The general characteristics of these training programs will be discussed (5.1.3.1), including those of the learners and trainers (5.1.3.2) as well is the content taught and the training systems used (5.1.3.3). Finally, the general effectiveness of these training programs, as formulated by the authors of the studies, in terms of the objectives pursued, will be considered (5.2.4).

5.1.3.1 Key Characteristics of Workplace Training

Two thirds of workplace training has the objective of reducing accidents, MSDs and/or pain. Changing behaviours is the second most important objective, but is quite far behind (27%) (Table 5.4). In rare cases, two objectives are determined (21%), but in general, studies only had one (67%). However, very few studies report the subject of the request at the origin of their workplace training intervention, i.e., the reason or reasons why the organization wanted a training program. The healthcare sector is by far the most interested in these training programs (60%). Twenty-eight percent of training programs were given in several workplaces during the same study. While that information is poorly reported, the size of organizations is variable, with a tendency for training to take place more often in large and very large organizations (23%) compared to small- and medium-sized organizations (10%). Slightly fewer than one in three studies provided information about the tasks for which training was requested (31%[7]). Fifty-six percent of the studies reported having completed a variety of phases before or at the same time as the training (phases reported: 27/48). In other words, 44% of the interventions were essentially limited to providing training, without any other additional activity within the organization. For the others, approximately one quarter of the studies (23%) involved preliminary investigations carried out in order to learn more about the environment and the handling tasks that were the subject of the training. One in five studies (19%) had a monitoring committee or group in place to support the intervention in the workplace and 31% transformed the work situation at the same time as the training took place. These transformations were most often trainee-oriented (e.g., handling assistance, personal protective equipment).

[6] The reader should make a clear distinction between the conclusions of the meta-analyses on the effectiveness of training and those of the authors of each study identified in these reviews. The former take into account the quality of the evaluation methodology, hence their harsher conclusions.

[7] Even when the information is provided, it remains general and often does not always provide a clear picture of the handling tasks: whether they are repetitive or varied; whether or not the loads are always similar; the characteristics of the worksites (e.g., handling heights, restricted spaces, stairs and slopes, circulation), etc.

Table 5.4 Key characteristics of workplace training programs

General characteristics	Total (n = 48[1])
Evaluation of the effectiveness of the training on[2]	
Reducing accidents/MSDs/pain	32 (67%)
Changing behaviours	13 (27%)
Improving physical capacities	4 (8%)
Improving knowledge	8 (17%)
Reducing health risk factors	6 (13%)
Other	7 (13%)
Number of simultaneously targeted objectives	
1	32 (67%)
2	10 (21%)
3 and more	6 (13%)
Information about the request	6 (13%)
Sector of activity	
Health/hospital	29 (60%)
Other sectors (construction, military, storage, etc.)	17 (35%)
Multiple	2 (4%)
Number of organizations involved	
1	33 (69%)
2 to 9	7 (15%)
10 or more	6 (13%)
Organization size	
Small	2 (4%)
Medium	3 (6%)
Large	5 (10%)
Very large	6 (13%)
Presence of information about work tasks	15 (31%)
Phases reported[1]	**27 (56%)**
Investigations	11 (23%)
Evaluation of risks	6 (13%)
Recommendations	6 (13%)
Diagnosis	2 (4%)
Changes	15 (31%)
Creation of a follow-up committee/work group	9 (19%)
Post-training follow-up	2 (4%)
Type of training[1]	
Skill transfer	19 (40%)
Physical conditioning	5 (10%)
Behaviour changes	15 (31%)
Raising awareness	7 (15%)

[1] *Three studies were removed from analysis, because they dealt solely with physical capacities*
[2] *Sub-categories are not mutually exclusive*

5.1.3.2 Profile of Trainees and Trainers

Overall, the studies provide more information about the trainees than about those providing the training, even though they remain incomplete (Table 5.5).

Table 5.5 Characteristics of trainees and trainers

Characteristics of…	**Total** (n = 48)
Trainees:	
Gender	
Mixed	21 (44%)
Male	3 (6%)
Female	6 (13%)
Average seniority	
0 to 5 years	2 (4%)
5 to 10 years	5 (10%)
10 years and more	1 (2%)
Health history	
Positive: with injury/MSD history	20 (42%)
Negative: no history	5 (10%)
Profile[1]	
Worker	47 (98%)
Student	2 (4%)
Manager	4 (8%)
Supervisor	6 (13%)
Occupation[1]	
Care staff	28 (58%)
Handler	14 (29%)
Housekeeping	4 (8%)
Other	11 (23%)
Handling experience	
Yes	26 (54%)
Mixed	3 (6%)
Trainers:	
Profile/Training[1]	
Ergonomist	6 (13%)
Occupational therapist/Physiotherapist	14 (29%)
Kinesiologist	1 (2%)
Worker	2 (4%)
Number of trainers involved	
1	13 (27%)
2 or more	9 (19%)

[1] *Sub-categories are not mutually exclusive*

The dominant profile of trainees in workplace training is that of workers (98%) in the healthcare field (58%) with previous handling experience (54%: few reported their years of seniority—8/48 or 16%). Some had a history of injuries/pain related to MSDs (42%). Women were significantly represented, although the groups were often made up of both men and women (mixed groups: 44%).

The trainers are healthcare professionals who are mainly physiotherapists or occupational therapists (29%), and, to a lesser extent, ergonomists (13%). In only two cases, training was provided internally by peers, i.e., by workers who were themselves handlers. More than one trainer was required in 19% of the cases, but several studies did not provide that information.

5.1.3.3 Content Taught and Training Systems Used

A significant proportion of training programs transmitted knowledge only (42%), with half of them (48%) combining knowledge and know-how[8] (Table 5.6). Although a fairly exhaustive list of topics was covered during the training sessions, the three main subjects discussed were, in order of importance: the safe handling technique (81%: see box below), the effects of handling tasks on health (60%) and the determinants (52%), including handling aids (44%), load characteristics (13%) and environmental characteristics (10%). One in four studies mentioned having adapted the content covered in training to the context in which the learners were working. Among these, 13% indicated having conducted preliminary workplace analyses and 4% used illustrations of actual working activities (photos or videos).

The safe handling technique: *straight back, bent knees*

The safe handling technique is defined in the studies as the method that should be used by handlers to protect themselves from injuries associated with handling. By looking at the studies (although this technique is not always described uniformly/identically), it is possible to list its main characteristics. Thus, the most commonly reported elements used to define this method are, in increasing order of mention in the studies, as follows:

- **Posture**: straight back, reduce torso flexion and avoid rotation;
- **Use the legs to lift**: bend the knees, use the large muscle masses of the legs to lift;
- **Lever arm**: keep the load close to the body, reduce the distance from the torso;
- **Balance**: be stable, maintain a wide support base;
- **Movements**: should be fluid and controlled, i.e., slow and smooth.

[8] All the variables used in the analyses are defined in Appendix 3.

Table 5.6 Content and structure of workplace training

Characteristics related to…	Total (n = 48)
Content:	
Knowledge transferred	
Knowledge only	20 (42%)
Knowledge and know-how	23 (48%)
Subject dealt with[1]	
Safe handling technique/good method	39 (81%)
Effects on health	29 (60%)
Determinants[1]	25 (52%)
Handling aid	21 (44%)
Load characteristics	6 (13%)
Environmental characteristics	5 (10%)
Work activity[1]	14 (29%)
Communication	3 (6%)
Work organization	6 (13%)
Knowledge transferred to coworkers	2 (4%)
Difficulties experienced at work	2 (4%)
Physical conditioning	14 (29%)
Pain management	9 (19%)
Injury and accident risk factors	13 (27%)
Prevention at work	7 (15%)
Management of OHS/standards/laws	4 (8%)
Transformations	4 (8%)
Individual aspects	3 (6%)
Suitability	
Adapted to the context[1]	12 (25%)
Including preliminary analyses	6 (13%)
Including illustrations	2 (4%)
Structure:	
Total duration	
< 1 day	13 (27%)
1 day	8 (17%)
≥ 2 days	7 (15%)
Location	
Only in class	18 (38%)
+ in the gym	3 (6%)
+ in a simulated environment	3 (6%)
+ in the field	13 (27%)
Only in the field	2 (4%)
Recommended educational approach	
Top-down/knowledge provider	29 (60%)
Mixed/transmits knowledge and promotes exchanges	8 (17%)
Procedure	
Theory only	9 (19%)
Practice only	4 (8%)
Theory and practice	28 (58%)
Feedback	16 (33%)

[1] *Sub-categories are not mutually exclusive*

Thirty-eight percent of training took place exclusively in classrooms, 4% uniquely in working situations and 27% in both locations, 6% in a gym/training room and 6% in a simulated environment. Almost two out of three training programs (60%) took a knowledge transfer approach, while a minority of them encouraged participation and exchanges between the trainer and the participants (17%). While most training programs added a practical workplace component (60%) to the theoretical component, one in five training programs was only theoretical, meaning that it involved no motor engagement from the learners, while 8% of them were solely practical in nature. Finally, one third of the studies included feedback during the training in order to guide trainees as they learned. However, the information available is uneven in terms of how the feedback was given, the nature of the information provided and what information was prioritized.

5.1.3.4 Effectiveness of Workplace Training

Table 5.7 presents results on the effectiveness of workplace training, as reported by the authors of the studies.

Table 5.7 General effectiveness of workplace training in terms of the objectives

Results	**Total** (n = 48)
General assessment formulated by the authors	
Positive assessment	22 (46%)
Mixed assessment	5 (10%)
No assessment	21 (44%)
Specific assessment for each objective (formulated by the authors)[1]	
Reduction in accidents/MSDs/pain	**32 (67%)**
Positive assessment	12 (25%)
No assessment	20 (42%)
Behaviour changes	**13 (27%)**
Positive assessment	9 (19%)
No assessment	4 (8%)
Improvement in physical capacities	**4 (8%)**
Positive assessment	4 (8%)
Knowledge enhancement	**8 (17%)**
Positive assessment	5 (10%)
Mixed assessment	1 (2%)
No assessment	2 (4%)
Reduction in health risk factors	**6 (13%)**
Positive assessment	4 (8%)
Mixed assessment	1 (2%)
No assessment	1 (2%)

Not mutually exclusive.

Almost one in two training programs (46%) had positive results and this proportion climbs to more than half if mixed/mitigated results (10%) are included. Forty-four percent had no effect. A distribution of the studies according to objectives shows that one study in four (25%) reports a reduction in accidents, pain and MSDs following the training program. This is a goal that most of the training programs wanted to reach and was that which was most associated with success, even if these provided no results in 42% of cases. These training programs also made it possible to reach other objectives, but with even more limited success.

5.2 Comparisons Based on Quality Criteria

Workplace training (n=48) was ranked so as to form more homogeneous groups and to compare their effectiveness (Table 5.8). Keep in mind that these groups were established according to the four quality criteria accepted in this study (for details about the criteria for each study, see Appendix 5). For example, all the training programs that reported having adapted their training content to the work context (criterion #1) were grouped together and compared with those for which the content has not been adapted: does the content have an influence on the effectiveness of training? Is enriched or adapted content more effective? This approach was used for the other three quality criteria[9].

Table 5.8 Effectiveness of workplace training (n=48) according to whether or not the quality criteria used in this study are taken into account in training

Results	Criterion 1		Criterion 2		Criterion 3		Criterion 4	
	Yes (n = 12)	No (N = 36)	Yes (n = 32)	No (N = 16)	Yes (n = 18)	No (N = 30)	Yes (n = 15)	No (n = 3
General assessment formulated by the authors								
Positive assessment	6 (50%)	16 (44%)	13 (41%)	9 (56%)	6 (33%)	16 (53%)	9 (60%)	13 (39
Mixed assessment	2 (17%)	3 (8%)	3 (9%)	2 (13%)	2 (11%)	3 (10%)	1 (7%)	4 (12
No assessment	4 (33%)	17 (47%)	16 (50%)	5 (31%)	10 (56%)	11 (37%)	5 (33%)	16 (48
n = 48	n = 12		n = 10		n = 6		n = 3	

Criterion 1: content adapted to the context; Criterion 2: motor engagement; Criterion 3: practice in a real or representative work environment; Criterion 4: changes to work situations

n — *Indicates the number of training programs that combine the criteria: 3 out of the 48 initial training programs have all the quality criteria*

Criterion #1: The impact of adapting training content to the context

First, we note the predominance of training programs that primarily focus on the safe technique (criterion #1: non-adapted content), which constitute three-quarters of workplace training programs (36 out of 48). It again appears that training programs in which the content is "adapted" also use the safe technique as the foundation of their approach. They aim to adapt the safe technique to the context or to improve it by adding more knowledge: there is therefore a difference between the content of the two groups, but also similarities. The results show that adapting the content to the context leads to a slight increase in effectiveness. While the authors

[9] In accordance with the theoretical framework for training presented in the methodology (sub-section 4.3.2.1), the authors' initial intention was to make comparisons between the training courses that include the first three criteria (n=6), or even all four (n=3), and the other courses. But as the grey cells show, the samples are too small to draw conclusions.

of both types of training report positive results in similar proportions (50% vs. 44%), adapting the content leads to a higher rate of mixed effects (17% vs. 8%). By combining the two, the impacts on effectiveness are clearer: 67% vs. 52%. It therefore appears that workplace training programs with content inspired by the safe technique but that attempt to go beyond it are slightly more effective than teaching the safe technique only.

Criterion #2: The impact of offering a motor engagement component

The proportion of training programs that include a motor engagement phase for participants in their training is twice as high as those that do not (32 vs. 16). It is the most characteristic quality criterion of training programs. However, adding a practical component does not appear to have resulted in better efficiency, in fact, the opposite has even been observed. The category without motor engagement leads to more positive results (56% vs. 41%): that difference is accentuated when we look again at the mixed results.

Criterion #3: The impact of practicing in a context representative of work

A similar phenomenon can be observed for practice in a context representative of work. Training programs that do not respect this criterion have better outcomes (53% vs. 33%). There is a close link between criteria #2 and #3: they are strongly related to motor engagement. Criterion #2 places greater emphasis on the need to repeat the movements to be learned, while #3 focuses on the environmental characteristics in which this practice will take place.

Criterion #4: The impact of transforming the work situation

The greatest gain in effectiveness is associated with the criterion not directly related to training. In almost one study in four, work situations were transformed (15 out of 48). Sixty percent of studies that implemented changes (e.g., purchase of handling assistance equipment) reported positive results, compared to 39% for training programs in which changes were not made at the same time. However, it should be noted that the main changes implemented were aimed primarily at handlers, and less at the determinants of the work situation that caused the restrictions that these workers must face (e.g., the weights of loads, space constraints).

The impact of combining the four criteria

Out of the 48 training programs selected, six (13%) simultaneously included the three criteria directly related to the quality of a training program (see subsection 4.3.2.1 for details). Three studies combined the four criteria, which represents slightly over 6% of all workplace training programs. With such a small sample, any attempt at comparison seems risky. However, it is the combination of criteria that best attests to high-quality training.

5.3 Methodological Quality of Evaluation Designs and Impacts on the Conclusions of the Meta-analyses

5.3.1 Quality of the Evaluation Methodology

The scores attributed to the methodological quality of evaluation designs by the authors of the HSE (Table 5.9) and Cochrane (Table 5.10) meta-analyses are presented first, and the classification details for the five reviews are found in Appendix 6. We first note the low proportion of studies that are assigned a high rating: keep in mind that these are the studies that had the greatest weight when the authors of the meta-analyses formulated their conclusions. From these ratings, three ascending groups[10] were created according to the quality of the evaluations (Table 5.11): high quality (n=7), good and average quality (n=12) and poor quality (n=12). In the following subsection, these groups were used to test the hypothesis of a selection bias in the meta-analyses.

Table 5.9 Classification of the methodological quality of articles by the reviews of the HSE group

Classification of methodological quality	Haslam et al., 2007 (n = 84)	Clemes et al., 2009 (n = 42[a])	Hogan et al., 2014 (n = 13)
High quality	7 (8%)	6 (14%)	3 (23%)
Other quality	37 (44%)	36 (86%)	10 (77%)
Good	7 (8%)	5 (12%)	6 (46%)
Average	13 (15%)	12 (29%)	3 (23%)
Poor	17 (20%)	19 (45%)	1 (8%)
Unassigned[b]	40 (48%)	-	-

[a.] 11 of 53 articles could not be identified.
Some articles were not the subject of an evaluation of their methodological quality (e.g., literature reviews, groups of experts, surveys carried out by questionnaires, audits, etc.)

Table 5.10 Classification of the methodological quality of the articles by the reviews of the Cochrane group

Classification of methodological quality	Martimo et al., 2007 (n = 12)	Verbeek et al., 2011 (n = 18)
Randomized controlled trial	5 (42%)	9 (50%)
High quality	2 (17%)	3 (17%)
Other quality	3 (25%)	6 (33%)
Poor	3 (25%)	6 (33%)
Cohort study	7 (58%)	9 (50%)
Other quality	7 (58%)	9 (50%)
High/good	7 (58%)	9 (50%)

10 We have excluded studies with contradictory classifications, as well as those that were unassigned. A total of 31 studies were used to construct the three groups.

5.3.2 Quality of the Evaluation Designs Compared to the Quality of the Training Evaluated

Table 5.11 compares the quality of the evaluation designs used in the studies (horizontal axis), as rated by the meta-analyses and grouped into the three ascending categories, and the quality criteria for training (vertical axis). The central hypothesis of this study is that there is a bias in the selection of training programs in the meta-analyses, which results in an inverse relationship between the rigour of the evaluation and the quality of the training actions considered. In other words, the training program to be evaluated had to be simple and uncomplicated to obtain the best possible evaluation (high-quality design).

This study's data appear to confirm that relationship. There is a tendency for the highest rated evaluation designs (i.e., the "high" category) to be assigned to training programs in which at least one of the four criteria attesting to the quality of a training program is less present, with the exception of criterion #4 concerning transformations.[11] For example, training content adapted to the context is found in 29% of the studies with a high evaluation methodology rating, while this proportion rises to 58% for a poor rating. For motor engagement, the percentages are 57% and 75%, respectively; for contextualized training, 29% vs. 50%. In addition, it should be noted that training programs in the "poor" category are those in which the simultaneous presence of several criteria combined is most prevalent: three studies included three criteria and two included all four (this is also the case for a study in the "other" category).

Table 5.11 Methodological quality of evaluation design versus training program quality

Quality criteria of a training program identified in this study	Methodological quality evaluated by the meta-analyses			Total (n = 31)
	High (n = 7)	Other[1] (n = 12)	Poor (n = 12)	
Criterion 1: ***Content adapted to context***	2 (29%)	1 (8%)	7 (58%)	10 (32%)
Criterion 2: ***Motor* engagement**	4 (57%)	8 (67%)	9 (75%)	21 (68%)
Aggregation of criteria 1 and 2[2]	1 (14%)	-	1 (8%)	2 (6%)
Criterion 3: ***Practice in a real or representative work context***	2 (29%)	4 (33%)	6 (50%)	12 (39%)
Aggregation of criteria 1, 2 and 3	-	-	3 (25%)	3 (10%)
Criterion 4: ***Transformations to work situations***	4 (57%)	5 (42%)	4 (33%)	13 (42%)
Aggregation of criteria 1, 2, 3 and 4	-	1 (8%)	2 (17%)	3 (10%)

1. Includes the studies with good and average methodological quality and the Cochrane group cohort studies.
2. Exclusively

11 Remember that this criterion does not concern training in the first place, but is intended to be a very useful complement in a global prevention approach.

6. DISCUSSION

This discussion is structured around the three objectives of this study. First, a standard portrait of handling training practices will be defined, as well as the training logic on which they are based (6.1). The conclusions of the meta-analyses with respect to the ineffectiveness of current training will be weighted: we shall see how the methodological quality of the evaluation designs has an impact on the conclusions. Second, the classification tests based on the identification of training quality criteria will be commented on (6.2). The fact that very few handling training programs have all of these quality criteria will be highlighted, as well as the resulting consequences. Finally, the authors will discuss a possible bias in the selection of training programs in the meta-analyses (6.3). In light of the current state of knowledge, the importance of evaluating training actions will be addressed, as well as the strict requirements associated with this evaluation, which seem ill-suited to identifying the effects of training approaches that have a greater level of complexity.

Finally, a critique of training focused mainly or even exclusively on mastering and adopting safe handling techniques will be presented (6.4). Particular emphasis will be placed on the limits of this approach in a context of self-regulation and on the idea that this is the main reason for the lack of effectiveness observed in handling training programs. The authors will defend the idea of the usefulness of, and even the need for, training to prevent handling risks, but will use a different approach that prioritizes the development of skills, the provision of more options or operating resources to handlers, and more flexibility.[12] It appears necessary to reframe the conclusions arrived at by the reviews on handling training: it is not so much handling training that is deficient but rather the type of training based on the predominance of the use of the safe handling technique.

To close the discussion, the limitations of this review will be discussed, in particular with regard to the main shortcomings in the material available to the research team, on which it was dependent to formulate conclusions (6.5). The representativeness of the material (the scientific literature), in terms of the realities of field practices in the provision of training, will also be addressed.

6.1 How Is Handling Taught?

In this section, the diversity of handling training practices and contexts will be discussed: the range of locations, clientele, occupations, etc. We will continue with what appears to be a paradox: the response to this variability is embodied in training content that promotes the use of universally applicable safety techniques. The philosophy underlying this training logic will be explained. The authors will draw their conclusions about the quality of handling training currently in use on the basis of this information.

---

[12] Definition (St-Vincent et al., 2014, p. 304): "Range of possible ways of doing things to adapt to the work situation: in other words, a worker's opportunity for self-regulation. The degree of leeway depends on (1) the person's characteristics, (2) the task requirements and the means available in the workplace."

6.1.1 Diversity of Training Practices and Contexts

The results show a diversity of training practices, both in terms of where people are trained, the objectives pursued, the educational methods used, the duration and the subjects covered. The objectives of handling training range from increasing the physical capacities of subjects (not necessarily handlers) in a controlled laboratory setting, to transferring skills in the workplace with workers of varying experience, and the acquisition of knowledge by future workers in training institutions. Of course, it is possible to highlight some key features. Among other things, the meta-analyses revealed that the healthcare sector appears to be a major consumer of these training programs; more training is provided in the workplace and in large institutions, with the main objective being to reduce accidents, pain, and MSDs. In these cases, the dominant educational paradigm is that of knowledge transmission. Nevertheless, handling training remains a vast and rather heterogeneous mosaic.

Unlike the Cochrane group reviews (Martimo et al., 2007; Verbeek et al., 2001), which were limited to workplace training, this study aimed to amass an inclusive sample so as to provide the most representative portrait of what is being evaluated in the handling training field. However, even when using this approach, which consists of limiting some of our analyses to workplace training, in order to describe training in more detail, the impression of broad diversity in practices remains. This heterogeneity is not in itself a surprise, since handling is carried out in all economic sectors, in all types of organizations, and the prevention issues related to handling are found in diverse disciplines with sometimes very different visions about what must be done to prevent risks in that field: the training approach is coloured by these disciplinary representations. However, a dominant vision is certainly that held by the biomedical field and the importance of individual factors in prevention.

6.1.2 Uniformity of Training Content

The heterogeneity of practices contrasts, however, with the surprising uniformity of content. For almost all the training programs, there is only one valid technique for handling a load correctly, but there are many ways to encourage its adoption by trainees. Even before beginning this study, the authors expected that the safe handling technique would be at the centre of learning. However, there was still some surprise at the omnipresence of this standardized work technique for which there seem to be few alternatives, except for a few suggestions that remain marginal. Despite some nuances, all of the training programs analyzed deal at one point or another with the need to keep the back straight and to bend the knees. Even the objective of improving physical capacities is linked to making it easier to adopt the safe technique when returning to the workplace. It is even used as such as a physical exercise for subjects, who are asked to perform repetitions of it.

The technique is thus presented as the only way to prevent handling injuries, although for decades studies have shown that it has limitations (Brown, 1973; Ayoub, 1982; Garg and Saxena, 1985). Those studies discuss the intrinsic limitations of the technique: high energetic expenditure, added strain on the knees, precarious balance, etc. More recent studies provide additional arguments that put the usefulness of the safe technique into perspective (Lortie, 2012; Denis et al., 2007). These call into question the how compatible the technique is with actual work contexts: it is often difficult to use because of the volume of loads or the demands of a rapid pace. However, over and above the criticisms that may be levelled about safe techniques, one fact remains: it is impossible for a single technique to be the only solution for all the variable situations that characterize handling activities. The reasoning underlying all of these training programs is that trainees and their behaviour must be the target of teaching. The limits of this paradigm are discussed in the following subsection.

6.1.3 A Training Logic Centred on Individuals, to the Detriment of Their Interactions with the Environment

This discussion will summarize the arguments against this technique, but will also identify its advantages. For now, the following hypothesis is explored: the emphasis placed on the use of this technique arises from a preventive view of handling in which individual behaviour is the origin of the injuries observed. This narrow focus on the safe technique is part of an overall training logic in which people and their behaviours are the target of the educational goals. Learners are asked to abandon their "bad work habits," to change their behaviour in favour of the systematic application of a standardized technique.

Little or no attention seems to be paid to the contexts in which these people work and their interactions with dynamic and changing environments, which are often imperfect in terms of handling conditions, and the associated demands of **self-regulation.**[13] Neither the learners' vocational experience nor their knowledge of work situations, in which some have worked for many years (Authier and Lortie, 1997), are really taken into account. To the contrary, experience seems to go hand-in-hand with the development of bad habits over time.

In conformance with deep-seated reasoning in some spheres of research, here we have a direct causal link, in which a cause (the "inadequate" behaviour of someone), leads to an effect in the form of an MSD. This link could prompt a seemingly simple solution: change the behaviour to one considered more compliant, this time considered as the "right way to work," to be applied at all times. This point of view goes well beyond the universe of research and appears to be deeply rooted, both among those requesting[14] training and many prevention experts. It leaves little room for alternative prevention strategies in which the handler is seen not as someone who simply executes the correct work method but as someone capable of making decisions based on the context in which he or she is working. According to that line of reasoning, individuals and their environment constitute a dynamic system, which is changeable and often unforeseeable, and which they must constantly regulate, in that it requires almost constant adjustments and cannot be known in advance. With this in mind, training takes on completely new dimensions, because the ubiquitous use of a single technique is incompatible with this need to self-regulate and adapt to variable work situations. However, what can we say about the quality of current handling training programs?

[13] According to St-Vincent et al. (2014, p. 304), "Self-regulation is a process of ongoing adaptation to varying work requirements and working conditions and to the worker's own personal variability. The worker has various strategies to make this adjustment. The goal of self-regulation is to maintain a balance between staying healthy and achieving work objectives."

[14] When requests for ergonomic internships are made for staff training, it is common for the employer to ask that training in the adoption of good work methods be provided, sometimes going so far as to question the competence of its employees to do their jobs.

6.1.4 Are Handling Training Programs of High Quality?

This question demands a nuanced response. In order to do so satisfactorily, it is necessary to position oneself with respect to training in relation to a theoretical framework,[15] as was the case in the section presenting the methodology. There is no single training approach that everyone agrees on: their quality is related to the theoretical positioning adopted by those wishing to respond to the question and to the arguments put forward to justify it. To reach a decision, four criteria attesting to the quality of training were determined to see whether or not they were taken into account in the training programs in our sample. The results show that approximately 10% of workplace training incorporates those four criteria.

According to proponents of a training approach focused on skills development, the information gathered to describe the training programs in the studies identified indicates that the level of quality of most of them is not good enough. The principal criticism is the training content, which, as we previously pointed out, is not adapted to the context of use, but rather aims to be universal and exportable from one place to another. A compromise was made in this study, because even the training programs that were classified as having an adaptable content did not totally respect the idea underlying this criterion: that skills used in context on a regular basis and that have a safety potential must also be considered. However, none of the studies identified met this criterion.

Nevertheless, all is not black and white. Despite the unfavourable conclusions of the meta-analyses, some training programs focusing on the safe technique do quite well in terms of effectiveness, as we will see in the following subsection.

6.2 Teaching Safe Techniques, but How Successfully?

The conclusions drawn by the meta-analyses regarding the effectiveness of the training programs identified, particularly with regard to the effects of the training, are surprising. However, the conclusions of the meta-analyses on the effectiveness of handling training programs (see p. 7) appear to leave little room for interpretation: the effects are minor, even nonexistent. Yet, consultation of the studies that led to this finding suggests a more nuanced view. The results show that one in two workplace training programs reports positive effects, with this proportion rising to more than 75% in the laboratory. Overall, more than half the studies (52%) reported positive post-training effects. How can this discrepancy be explained?

An interesting observation can be made: according to the meta-analyses, barely one in ten studies is considered to be of high quality in terms of the training program evaluation process. Despite the impressive number of studies devoted to handling training and its evaluation, the meta-analyses formulate their conclusions on the basis of only a limited number among them, because they give precedence to those that are based on evaluation designs recognized as being of high quality (e.g., with control groups, randomized trials, pre-/post-evaluations with longitudinal follow-up). The studies rated as good or average would also be used to draw the conclusions, but they would have little or no weight, while those with a poor design would have little or no chance of being included. This is a limitation that should be taken into account: not only is the sample small, but it is questionable as to whether it is representative of the best handling training programs available.

[15] This information was not presented in the results, but a minority of studies present the training logic (paradigm) that they use or that theoretically justifies their approach.

In fact, the data should be regarded cautiously. Between the mixed conclusions of the meta-analyses and the overall positive results reported by the authors of the studies identified, there should be something in between. The position of this research team is that the safe technique has the potential to prevent accidents, but it is certainly not (or no longer) the only one. Some would argue that an evaluation design that attests to the effectiveness of a training program is key, and if it is not sufficiently thorough, one cannot trust the conclusions drawn from it. But, at the same time, does the predominance given to these high-quality designs conceal another less evident reality, which is just as prejudicial: should decisions be made about the effectiveness of training programs based on a biased sample? Are the best training practices excluded because of methodological deficiencies that do not adequately assess their effects? Which approach should be given priority: the science of the evaluation and the scientific rigour of its design?[16] Or should the priority be on the science of education and its learning principles, which do not seem to lend themselves easily to the dictates of those who evaluate them?

6.3 Does the Challenge of Evaluating Training Programs Conceal a Bias?

Let us be clear: it is not a matter of questioning the logic of those whose objective is to thoroughly assess the effectiveness of interventions, and training is no exception. They must base themselves on high-quality data. At the same time, does this mean one must penalize or exclude approaches that, because of their great complexity, cannot be easily evaluated? Does this mean that which cannot be measured or which is more difficult to evaluate has less value?

As stated in the introduction, in a review of the literature on MSD prevention interventions (Denis et al., 2008), a similar phenomenon was observed: few studies combined a state-of-the-art approach with a proven evaluation phase. This is a paradox that can become a dilemma. On the one hand, if one wishes to conduct evaluations with the thoroughness required, it seems that a compromise must be made about the quality of the training on the effects that one wants to evaluate. On the other hand, if the quality of the training is good, there is always a double risk: that of being almost certain that it cannot be adequately assessed and that therefore, it will then be ignored by the meta-analyses that seek to judge the effectiveness of training in a given field.

This situation is of interest to us for two reasons. Often, the first question asked about the expected results following training by those requesting it is "what's in it for us?" This question, from the viewpoint of a researcher-practitioner, is very important in justifying the actions that will be taken. The development of new evaluation approaches compatible with the contingencies of the field and/or with the particular features of a higher-level training activity is desirable. The second concern relates to the conclusions drawn by the meta-analyses conducted on handling training practices. For those who do not have the opportunity to analyze the various handling training offers, the conclusions of the reviews can be interpreted as saying that handling training has negligible effects. That conclusion is wrong: in light of this study's results, it would be more accurate to say that handling training limited to the transmission and adoption of safe handling techniques provides very few effects in terms of prevention. The distinction is important because, in the first case, the normal reflex is to abandon the training avenue as a preventive activity in handling, or to limit investment in resources. Alternatively, the reaction could be to

[16] The authors wish to point out that scientific rigour is not limited to evaluative research, experimental design or the biomedical universe. Each methodological approach, whether inspired by the pure sciences, social sciences or the humanities, has its own criteria for quality, rigour and accuracy. However, it must be noted that rigour is more often associated with quantitative approaches, in particular, those that use statistical processing methods.

change training practices so that they are more in line with the actual activity of handlers. This is the avenue we will pursue.

6.4 A Fresh View of Handling Training

6.4.1 The Safe Technique: Its Limits and Its Advantages

First, let us briefly discuss the origin of the safe technique and the reasons why it is believed to prevent injuries associated with handling. Primarily developed from laboratory studies in the field of biomechanics (Sedgwick and Gormley, 1998), this technique has the goal of reducing mechanical overload on the spine, especially in the lumbar region. Mechanical stress is recognized as a principal cause of back pain, which is why so much importance is attached to it. That is why "safe" handling techniques are encouraged (Authier and Lortie, 1995). The emphasis is on the lifting phase, because it is then that the maximal loading of the back takes place (Plamondon et al., 2014). Biomechanics explains the distribution and intensity of the body's efforts to respond to external demands. In that sense, and in the context of training, the resultant recommendations consist in suggesting standard procedures to protect the back from overload: being close to and facing the load, keeping the back straight and the knees bent, ensuring one's balance, moving the load slowly without jerking, etc. The goal of these techniques is to ensure that the load is distributed evenly over the spine and to facilitate the type of stress that the spine is better able to withstand (i.e., compressive forces), without placing too much stress on passive structures such as ligaments.

Even so, almost 30 years ago, Wax et al. published an article that reported arguments against the safe technique, which was particularly criticized in the United States at that time (Wax et al., 1987). The three arguments that were presented are still relevant today:

- The use of this technique requires significant energetic expenditure because it requires lifting one's own body weight in addition to the load every time. The displacement of the body's centre of gravity is then longer and goes against gravity;

- Work situations, such as an obstructed space (which restricts foot positioning), and the volume and/or shape of loads that do not easily fit into the base of support make it difficult to use. Other authors report that this technique is incompatible with the production objectives of organizations, because it requires more time and therefore slows down the work pace (Garg and Saxena, 1985);

- In the case of heavy loads, the technique could cause knee joint injuries, due in part to intense stress on the quadriceps. Lortie et al. (1993) also reported difficulties in keeping the body balanced while in a crouched position, both because the individual is on tiptoes and because when the knees are bent they make it difficult to bring the load closer to the body.

In contrast, the study by Wax et al. also demonstrated the usefulness of the "straight back, bent knees" technique under specific and well-defined conditions: lifting, by trained subjects, of a 25 kg[17] load, placed on the ground, with handles at the top of it,[18] at a rate of three lifts per minute

[17] Constitutes the weight limit for a man under "ideal" conditions (ISO 11228-1, 2003: lifting from the ground not being an ideal condition), with a threshold of 15 kg for women. This standard did not exist at the time the Wax et al. study was written.

for 20 minutes. Compared to the so-called spontaneous or free technique, the authors observed a reduction in biomechanical stress without additional energy expenditures.

Based on these findings, what can we conclude about the use of the safe handling technique? As is generally the case with any work technique, the list of advantages and disadvantages of a given way of doing things is always assessed according to the configurations of the work situation. To summarize:

- The list of advantages of a given work method overlaps with those of the disadvantages. There will always be advantages and disadvantages to a particular way of doing things;

- Therefore, no technique can be described as being **good in absolute terms**. However, there may be a **better way of doing something in a given workplace setting**: if there are more advantages than disadvantages to a technique, its use is justified;

- The safe handling technique is no exception to this rule. Its use is justified under certain conditions, as shown by Wax et al. (1987). The advantages in terms of stress distribution in the lower back are likely to outweigh the disadvantages associated with increasing energetic stress or slowing down the pace. The worker then makes a compromise.

Thus, in a regularly changing context, there is no such thing as identical execution or repetition of tasks: there are instead adaptations of varying magnitude, based on decision-making and judgments. However, we may ask the following question: approximately what percentage of the situations experienced by handlers during their work shift justifies using the safe handling technique? There is no precise answer to this question, but one could argue that the percentage is much lower than what people generally believe. As evidence of this, in a study carried out in the retail trade, heavy loads sitting on the ground represent less than 20% of handling tasks in a work shift (n=452 handling activities observed: St-Vincent et al., 2004). However, to deal with most situations (more than 80%) that do not justify use of the safe technique, what other ways of doing things are taught in handling training courses? None, according to the results of this review. This fact not only deprives handlers of alternative ways of performing their tasks, but it also does not provide them with the opportunity to develop their judgment and problem-solving skills, because they are seen simply as those who execute an action.

It is difficult to explain with certainty the reasons why the safe technique takes up so much space, but here are some arguments:

- **The variability of handling situations is underestimated.** Once one becomes familiar with the handling environment, one realizes that handling tasks are extremely diverse. They require handling varied loads, in diverse and changeable work environments (e.g., deliveries to a wide variety of clients, preparation of various orders, weather conditions when working outdoors). However, variability means adaptation and self-regulation;

- **Some risks are also underestimated.** As the back is the part of the body that is most affected, and overexertion is the most often reported causal agent in accident

[18] For Plamondon et al. (2012), placing a load 30 cm off the ground significantly reduces the external movements at the L5/S1 joint. The boxes used in the Wax et al. experiment were 0.25 m (25 cm) high, which undoubtedly had an impact on the observed (absolute) lumbar stresses.

databases,[19] research has been highly focused on that area. However, we know that the risks associated with handling are diverse and can vary greatly depending on the context (Lortie et al., 1996; Lortie and Pelletier, 1996; Lortie, 2003). A fundamental aspect concerns the frequency of handling and the possible link with the development of general and/or local fatigue. Although studies have documented this type of strain (Gallagher et al., 2007; Wilson et al., 2006; Granata et al., 2004), prevention is primarily concentrated on establishing threshold values (e.g., maximum handling frequency, daily tonnage), but, unlike the findings of studies on lumbar overexertion, no training recommendations have been formulated.

- **Handling training is first and foremost an issue related to individuals and the dominant approach is to show handlers what they must do.** This last reason appears even more plausible, especially since a number of indicators from this systematic review point in that direction:

  - The studies remain silent about the specific handling tasks for which training is developed, as if knowing more about these tasks has no value in adjusting training content to their particular characteristics. The lack of descriptions of the tasks involved is not a choice, but rather a symptom of a lack of interest in them and an underestimation of their role. In the same vein, the few references to work-related aspects in training content reinforce this observation;

  - Few studies report having conducted preliminary analyses to adapt and contextualize the knowledge and skills to be taught. This means that the content transmitted may be exported from one workplace to another, regardless of whether the worker is a mover or an employee in the construction sector;

  - Very often, the training program is the only action carried out in context. No other action is carried out in the workplace. However, when work situations are transformed in parallel with the training, which was the case for one study in five, the actions were mainly concerned with handling aids or personal protective equipment. These actions therefore relate primarily to the trainees, and not to the context.

Thus, the learners appear to be isolated from the context in which they work. They are not perceived as interacting with a dynamic environment that requires adaptation.[20] On the contrary, in order to determine training content, it is essential to fully understand these workers' activities and the environment in which they work. The compromises they must make to deal with often imperfect work situations must also be taken into consideration. It is also the ideal opportunity to understand how the context impedes the use of certain more appropriate work techniques and to make changes. The worst contexts, those that are the most restrictive for handlers, are those in which they are always forced use the safe technique.

---

[19] The studies show that many risks associated with handling are not taken into account in databases and there are many classification difficulties (Manning et al., 1984; 1988).

[20] It should be noted that workers not only adapt themselves to their context, they will sometimes transform and shape it. The activity is not solely reactive; it is also proactive, in the sense that it enables the outcome to be anticipated in order for workers to prepare and protect themselves. While this nuance is important, the authors preferred to highlight the idea of interaction between workers and their context, without going too far into the subtleties it may imply. This reservation is justified by the current gap between the idea of interaction and the training practices reported in the articles identified.

6.4.2 How Can Training Be Improved?

The field studies conducted in various handling contexts help us understand the gap between the techniques recommended in training and what is actually being done in the workplace (for a more detailed summary and portrait, see Denis et al., 2013). These studies paved the way for the three approaches suggested here to improve training programs[21].

a. Improving content:

a1. Enriching the "gestural vocabulary" of handlers: given the variability of the conditions in which handling takes place and the associated risks, there cannot be only one way of performing handling tasks without injury. The safe techniques currently taught must be among the skills available, but they must not become the entirety of the gestural repertoire of handlers. It has become clear that a handling training program must be based on handlers' know-how, especially those with many years of experience;

a2. Learning to choose an appropriate action: know-how is important, but it is not enough to prevent handling injuries. The key to prevention is in choosing an action adapted to the situation: it is essential to ensure that the advantages of that action remain greater than the disadvantages, depending on the context. The ability of handlers to analyze the situation and to find a suitable solution is crucial in protecting against the risk of injury, while enabling them to meet the production objectives imposed on them;

a3. Learning to organize one's work: beyond handling loads one by one, there is the need to introduce work-planning principles. The demands of production (i.e., daily tonnage) and the organizational methods established by the employer (e.g., teamwork, stability of assignments, time organization) will influence how handlers organize their work and plan their tasks so as not to have to (re)handle the same loads unnecessarily. In interviews with delivery workers, where they were asked to explain what was the most important, all mentioned aspects of knowledge related to the work context: knowledge of clients, layouts, streets (Lortie, 1982). This helps them better plan their work and choose the most appropriate methods of transport.

b. Establishing Conditions Conducive to Learning

Once the training content has been determined, it is necessary to focus on the elements of the educational system. Organizing the learning path for new workers is a challenge for the trainer and the organization, which must maintain its level of productivity and its competitiveness.

b1. Encourage motor engagement: handling remains above all a physical activity that requires motor learning: motor skills predominate. In fact, knowledge is encoded by the body. Training must therefore provide sufficient practice time. Knowing how to make the right decisions based on the context is not easily learned either. Practice is all the more important when several different techniques are being taught and the aim is to show learners how to choose the best one according to the work situation configurations. The greater the variability, the longer it will take to learn.

[21] These proposals were formulated in the framework of development of a new handling training program. They are based both on our studies of handlers and on the limitations observed in handling training. The results of this review confirm our empirical observations of training dynamics, in addition to adding some new elements to support them.

The results of this review, however, show that motor practice did not seem to have an effect in terms of prevention. On the one hand, it could be assumed that the duration of motor practice was not long enough to produce effects. Too few studies reported in sufficient detail the nature and duration of these practice periods to be able to determine the benefits that may result. On the other hand, and more realistically, it is logical to think that devoting time to practice a single handling technique, which, as we have seen, has many limits in its application, will have few effects in preventing MSDs: what is the point of practicing a skill with extremely limited applicability? It is therefore not the value of motor engagement that must be questioned, but rather, what is being practiced. We remain convinced of the value of encouraging motor engagement.

b2. Practice using the real working situation, but adapt the level of difficulty: It is suggested that cognitive processes change according to the goals trainees set for themselves, for example between learning goals and performance goals (Lauzier and Denis, 2016). Thus, whether it is during training or immediately after the return to work, trainees should not be placed in regular production conditions, nor should they be asked to perform their tasks alone. The working conditions should gradually be adapted to the level of competence and adjusted to trainees' progress.

c. Transforming Handling Situations:

c1. Change the conditions in which handling takes place for the better: This review shows that changing the work situation in parallel with training can improve effectiveness. The meta-analyses come to the same conclusion. Although transformations that directly affect trainees remain relevant, changes to work situations should also be encouraged. Efforts to modify the work environment also send a message to trainees that prevention does not rest solely their shoulders. Moreover, these transformations may also make it easier to use the skills learned in training.

To summarize, Table 6.1 compares what is currently being done in handling training and the changes that would be useful to implement.

Table 6.1 Comparison between current and desired handling training programs

WHAT IS DONE CURRENTLY	*VS.*	*WHAT TO AIM FOR*
Focused on transmitting prescribed methods to be applied all the time		Focus on developing skills where the work situation becomes central
Emphasis on the lifting phase		Consider all handling phases: lifting/carrying, depositing
Centred on preventing overexertion of the lumbar spine		Take all risks into account: overexertion and accumulation, fatigue and incidents
Turnkey training, generic		Contextualized, environment-specific training
Focused on the physical dimension and handling loads one by one		Consider the cognitive component through analysis of the work and ability to organize
Short course with an emphasis on theory, classroom lecture		Longer training program, with the emphasis on practice; training and action at the workstation(s)
Prevention is based on training and supporting the worker		A preventive approach that includes training and action on the other determinants (working conditions)
"Expert" type approach where the trainer is the source of knowledge		A participatory approach where workers' expertise is put to good use

6.5 The Limitations of this Study

There are three main limitations to this study. The conclusions reported by the authors of the studies were used to estimate the effects arising from training programs and the various cross-tabulation exercises undertaken. However, this does not take the quality of their evaluation design into account, as did the meta-analyses (this information is reported in the tables where relevant). As a result, their effectiveness is possibly overestimated. Examining the quality of studies would not have been an easy task to achieve, because the size of the samples would have been too small to draw conclusions.

On the other hand, the research team was dependent on the amount and, above all, the quality of the information in the articles, which were generally rather lacking in detail. For example, an author may have mentioned that the trainer was encouraged to provide feedback to the trainees during the training, without providing any details. Thus, no data were sufficiently comprehensive to be able to comment on the nature of the feedback, its frequency, quality, etc. All the articles that mentioned having used feedback were therefore processed identically, when there may have been notable differences.

It is legitimate to ask to what extent training programs that are the subject of scientific publications (with all the limitations associated with that exercise) reflect the practices of those involved in the field. Of course, some practitioners do things in differently, and their activities are not necessarily reflected in the writings of the scientific community. However, it is clear that these initiatives remain marginal compared to the popularity of the "straight back, bent knees" technique, which remains the benchmark.

7. CONCLUSION

Two key conclusions can be reached. The first is that training focused mainly on learning and adopting the safe handling technique has little impact in terms of prevention. Although more than 50% of studies reported positive effects, this percentage is low when one considers the quality of the evaluation designs used. In that approach, the focus is on individuals and their behaviours, without any connection with the work context. This aspect of training explains its ineffectiveness in preventing risks. The content must be reviewed to provide more operational possibilities to handlers to help them adapt to their changeable working situations. Training should not target the individual, but rather the interactions between the individual and the work context. To achieve this, the trainer must therefore be familiar with or at least understand the context before providing training, so as to adapt the teaching and to contextualize it. The safe technique should be one of the skills taught, because we know that it has mechanical advantages in specific work settings, but should not be the dominant one. The problem is not so much the safe technique, but the emphasis placed on its use, without sufficient consideration of the work context.

The second conclusion concerns the difficulties inherent in the evaluation of training programs. Few of the studies in the sample analyzed are based on an evaluation methodology that meets the relevant quality standards. There is a lack of arguments justifying training actions. However, it appears that this evaluation effort is sometimes to the detriment of the subject being evaluated. Evaluation methods seem to be better adapted to assess "simple" training, but is that type of training the most appropriate to prevent risks and prepare workers for their tasks? And if a training program can play such a role, which should its characteristics be? In that regard, this study provides a partial response to these groups of workers, who are among the most affected by musculoskeletal disorders.

BIBLIOGRAPHY

Allaire, M. and Ricard, G. (2007). *Statistiques sur les affections vertébrales 2003-2006*. Commission de la santé et de la sécurité au travail (CSST) du Québec, Direction de la statistique et de la gestion de l'information. Service indicateurs de gestion, analyses et études, pp. 1-29.

ASCC (2007). *National code of practice for the prevention of musculoskeletal disorders from performing manual tasks at work*. ASCC, Canberra, Australia, 133 p.

Authier, M. and Lortie, M. (1997). How do expert handlers choose a work technique? 13th Congress of the IEA, Tampere, Finland.

Authier, M. and Lortie, M. (1995). Y a-t-il une bonne méthode de manutention? Le point de vue de manutentionnaires experts. *Travail et santé*, 11, S-2-S-5.

Ayoub, M. M. (1982). The manual lifting problem: the illusive solution. *Journal of Occupational Accidents*, 4, 1-23.

Bernier, S., Decornois, I., Freyria, A.M., Lomberget, J.P., Lucas, C. and Sanglerat, M.B. (2003). *La prévention des risques liés aux manutentions manuelles et mécaniques - S'organiser et apprendre à bien porter pour mieux se porter*. CNRS. Meudon, France.

Bigos, S. J., Holland, J., Holland, C., Webster, J. S., Battie, M. and Malmgren, J. A. (2009). High-quality controlled trials on preventing episodes of back problems: systematic literature review in working-age adults. *The Spine Journal*, 9, 2, 147-168.

Blue, C L. (1996). Preventing back injury among nurses. *Orthopaedic Nursing*, 15, 6, 9-22.

Brown, J. R. (1973). Lifting as an industrial hazard. *Am. Ind. Hyg. Assoc. J.*, 34, 292-297.

Chardon, O. and Estrade, M.-A. (2007). *Les Métiers en 2015*. Joint publication of the Centre d'analyse stratégique and DARES, Collection "Qualifications et Prospectives." www.strategie.gouv.fr

Clarke, M. and Oxman, A.D. (2002). *Cochrane Reviewers' Handbook 4.1.5 [updated April 2002]*. Oxford: Update Software. Updated quarterly.

Claudon, L., Cuny-Guerrier, A., Desbrosses, K., Gilles, M.A. and Savescu, A. (2016). 9th International Scientific Conference on the Prevention of Work-Related Musculoskeletal Disorders, PREMUS 2016 (Toronto, Canada, June 20-23, 2016). *Références en santé au travail,* 148, 91-98.

Clemes, S.A., Haslam, C.O. and Haslam, R.A. (2009). What constitutes effective manual handling training? A systematic review. *Occupational Medicine*, 60, 101-107.

Denis, D., St-Vincent, M., Lortie, M., Gonella, M. and Dion, M-H. (2011). *Analyse des activités de manutention de journaliers d'une grande municipalité québécoise: un outil pour composer avec le caractère changeant de la manutention*. Montréal, IRSST, 74 pages.

Denis, D., Lortie, M., St-Vincent, M., Gonella, M., Plamondon, A., Delisle, A. and Tardif, J. (2013). *Participatory Training in Manual Handling Theoretical Foundations and Proposed Approach*. Montréal, IRSST, 142 p.

Denis, D., St-Vincent, M., Imbeau, D., Jetté, C. and Nastasia, I. (2008). Intervention practices in musculoskeletal disorder prevention: A critical literature review. *Applied Ergonomics*, 39, 1, 1-14.

Denis, D., St-Vincent, M., Gonella, M., Couturier, F. and Trudeau, R. (2007). *Analyse des stratégies de manutention chez des éboueurs au Québec - Pistes de réflexions pour une formation à la manutention plus adaptée*. Montréal, IRSST, R-527.

Downs, S.H. and Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomized and non-randomized studies of health care interventions. *Journal of Epidemiology Community Health*, 52, 377-384.

Furlan, A.D., Pennick, V., Bombardier, C., van Tulder, M. and the editorial board of the Cochrane Back Review Group (2009). Updated method guidelines for systematic reviews in the Cochrane Back Review Group. *Spine*, 34, 18, 1929-1941.

Gallagher, S., Marras, W.S., Litsky, A.S., Burr, D., Landoll, J. and Matkovic, V. (2007). A comparison of fatigue failure responses of old versus middle-aged lumbar motion segments in simulated flexed lifting. *Spine*, 32, 1832-1839.

Garg, A. and Saxena, U. (1985). Physiological stresses in warehouse operations with special reference to lifting technique and gender: a case study. *Am. Ind. Hyg. Assoc. J.*, 46, 2, 53-59.

Granata, K.P., Slota, G.P. and Wilson, S. E. (2004). Influence of Fatigue in Neuromuscular Control of Spinal Stability. *Human Factors*, 46, 81-91.

Guérin, F., Laville, A., Daniellou, F., Durrafourg, J. and Kerguelen, A. (1997). *Comprendre le travail pour le transformer, la pratique de l'ergonomie*. ANACT, Collection Outils et Méthodes, 2nd edition.

Haslam, C., Clemes, S., McDermott, H., Shaw, K., Williams, C. and Haslam, R. (2007). *Manual handling training – Investigation of current practices and development of guidelines*. Loughborough: HSE, RR583.

Hermans, V., De Preter, G. and Verschueren, T. (2012). Training in manual material handling: what is going on in the field? *Work*, 41, 588-591.

Higgins, J. and Green, S., eds. (2008). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.0 [updated February 2008]*. The Cochrane Collaboration.

Hogan, D.A.M., Greiner, B.A. and O'Sullivan, L. (2014). The effect of manual handling training on achieving training transfer, employee's behaviour change and subsequent reduction of work-related musculoskeletal disorders: a systematic review. *Ergonomics*, 57, 1, 93-107. ISO 11228-1 (2003). Ergonomics—Manual handling—Part 1: lifting and carrying.

Jonnaert, P. (2009). *Compétence et socioconstructivisme: un cadre théorique*. Brussels: De Boeck Université (2nd ed.).

Kirkpatrick, D.L. (1994). *Evaluating Training Programs: The Four Levels*. San Francisco, CA: Berret-Koehler.

Lauzier, M. and Denis, D. (2016). *Accroître le transfert des apprentissages: vers de nouvelles connaissances, pratiques et expériences*. Presses de l'Université du Québec, Sainte-Foy, Canada.

Le Boterf, G. (2002). *Développer la compétence des professionnels. Construire les parcours de professionnalisation*. 4th edition, Éditions d'organisation, Paris, France.

Ledoux, E. and Laberge, M. (2006). *Bilan et perspectives de recherche sur la SST des jeunes travailleurs*. Report R-481, Montréal, IRSST, 80 pages.

Lortie, M. (2012). Manutention: prise d'information et décision d'action. *Le travail humain,* 65, 193-216.

Lortie M. (2003). L'analyse des risques associés aux activités de manutention à caractère variable. 38th SELF conference, Paris, September 24-26, 413-419.

Lortie, M. and Pelletier, R. (1996). Incidents in manual handling activities. *Safety Science*, 21, 223-237.

Lortie, M., Lamonde, F., Collinge, C. and Tellier, C. (1996). Analyse des accidents associés au travail de manutentionnaires sur les quais dans le secteur transport. *Travail humain*, 59, 2, 187-205.

Lortie, M., Baril-Gingras, G. and Authier, M. (1993). Manutention et risque? *Performances Humaines & Techniques*, 63, 23-28.

Lortie, M. (1982). Choix des variables dans la conception d'un système de livraison. 18th SELF conference, 203-204.

Mairiaux, P., Demaret, J.-P., Freyens, S., Masset, D., Vandoorne, C. and Vandecan, J. (1998). *Guide pour évaluer et prévenir les risques - Manutentions manuelles*. Commissariat Général à la Promotion du Travail, Brussels, Belgique.

Manning, D.P., Mitchell, R.G. and Blanchfield, P.L. (1984). Body movements and events contributing to accidental and non-accidental back injuries. *Spine*, 9, 7, 734-739.

Manning, D.P., Ayers, I., Jones, C., Bruce, M. and Cohen, K. (1988). The incident of underfoot accidents during 1985 in a working population of 10,000 Merseyside people. *Journal of Occupational Accidents*, 10, 121-130.

Martimo, K. P., Verbeek, J., Karppinen, J., Furlan, A. D., Kuijer, P. P. F. M., Viikari-Juntura, E., Takala, E. P. and Jauhiainen, M. (2007). *Manual material handling advice and assistive devices for preventing and treating back pain in workers (Review).* John Wiley & Sons.

Plamondon, A., Delisle, A., Bellefeuille, S., Denis, D., Gagnon, D. and Larivière, C. (2014). Lifting strategies of expert and novice workers during a repetitive palletizing task. *Applied Ergonomics*, 45, 471-481.

Plamondon, A., Larivière, C., Delisle, A., Denis, D. and Gagnon, D. (2012). Relative importance of expertise, lifting height and weight lifted on posture and lumbar external loading during a transfer task in manual material handling. *Ergonomics*, 55, 1, 87-102.

Provencher, J., Barbeau, A., Lamarche, D., Robitaille, D. (2011). *Statistiques sur les affections vertébrales 2007-2010-CSST*. Service de la statistique et direction de la comptabilité et de la gestion de l'information, Commission de la santé et de la sécurité du travail du Québec (CSST), 28 p.

Rivard, P. and Lauzier, M. (2013). *Les modèles et principes d'apprentissage dans la gestion de la formation et du développement des ressources humaines: pour préserver et accroître le capital compétence de l'organisation*. Presses de l'Université du Québec, 2nd edition, 67-72.

Rodrick, D. and Karwowski, W. (2006). Manual Materials Handling (Chapter 30). In *Handbook of Human Factors and Ergonomics* (third edition), Edited by Gavriel Salvendy, published by John Wiley & Sons, Inc., Hoboken, New Jersey, 818-854.

Sedgwick, A.W. and Gormley, J.T. (1998). Training for lifting: an unresolved ergonomic issue? *Applied Ergonomics*, 29, 395-398.

Schmidt, R.A. (1988). *Motor Control and Learning: A Behavioral Emphasis*. 2nd ed. Champaign, IL: Human Kinetics.

Slim, K., Nini, E., Forestier, D., Kwiatkowski, F., Panis, Y. et and Chipponi, J. (2003). Methodological index for non-randomized studies (MINORS): development and validation of a new instrument. *ANZ journal of surgery*, 73, 9, 712-716.

St-Vincent, M., Vézina, N., Bellemare, M., Denis, D., Ledoux, É. and Imbeau, D. (2014). *Ergonomic Intervention*. Montréal, IRSST, 336 p.

St-Vincent, M., Laberge, M., Denis, D., Richard, M-C., Imbeau, D., Delisle, A. and Dufour, B. (2004). *Les principaux déterminants de l'activité de manutention dans un magasin-entrepôt de grande surface*. Montréal, IRSST, R-365, 77 p.

Tveito, T. H., Hysing, M. and Eriksen, H. R. (2004). Low back pain interventions at the workplace: a systematic literature review. *Occupational medicine*, 54, 1, 3-13.

van Tulder, M.W., Furlan, A., Bombardier, C., Bouter, L. and the editorial board of the Cochrane Collaboration Back Review Group (2003). Updated method guidelines for systematic review in the Cochrane Collaboration Back Review Group. *Spine*, 28, 12, 1290-1299.

Verbeek, J.H., Martimo, K.P., Karppinen, J., Kuijer, P.P.F.M., Viikari-Juntura, E. and Takala, E.P. (2011). *Manual material handling advice and assistive devices for preventing and treating back pain in workers (Review)*. John Wiley & Sons.

Wax, C., Flenghi, D. and Meyer, JP. (1987). Comparaison de deux techniques de lever de charges. Analyse biomécanique et coût physiologiques. *Le Travail Humain*, 50, 335-345.

Wilson, E.L., Madigan, M.L., Davidson, B.S. and Nussbaum, M.A. (2006). Postural strategy changes with fatigue of the lumbar extensor muscles. *Gait & Posture*, 23, 348-354.

Appendix 1 – Conclusions of Five Meta-analyses

Reviews	Conclusions	Results
	HSE Group	
Haslam et al. (2007)	Weak evidence of the effectiveness of handling training. No transfer of learning in the workplace (considerable evidence) Strength and flexibility conditioning appear to be beneficial. However, the long-term effects of these activities on reducing risks in the workplace must be studied. No evidence of the effectiveness of "back school" type training on the prevention of low back pain. Evidence of the effectiveness of ergonomic training/activities[1] on reducing the risks of accidents related to handling.	Mixed
Clemes et al. (2009)	Handling training is not effective in reducing back pain and injury. No transfer of learning in the workplace (considerable evidence) Strength and flexibility conditioning have potential, but the long-term effects of these activities in reducing MSDs must be studied.	Mixed
	Cochrane Group	
Martimo et al. (2007)	Handling training is not effective in reducing back pain and injury, with or without the addition of handling aids. (Moderate to weak evidence)	Nil
Verbeek et al. (2011)	Handling training is not effective in reducing back pain and injury, with or without the addition of handling aids. (Moderate to very weak evidence).	Nil
	Other	
Hogan et al. (2014)	Handling training does not appear to be effective in reducing work-related MSDs. Employees reported that they had acquired knowledge (not at the level expected), but it did not always lead to behaviour change.	Nil

1. E.g.: evaluation of risks, modifications to equipment or tasks, etc.

Appendix 2 – Meta-analysis Inclusion Criteria

Meta-analysis Inclusion Criteria

HSE Group

Haslam et al., 2007 (n = 84)

Year of publication: 1980 to 2006
Methodology: peer-reviewed article, health and safety agency report and conference proceedings
Activity: empirical study conducted in the laboratory or workplace, handling training or physical conditioning, survey by questionnaire, audit on the effectiveness of handling training, literature review on the effectiveness of handling training
Study objective: to evaluate the effectiveness of material handling training or the impact of physical conditioning on improving manual handling performance

Clemes et al., 2009 (n = 53)

Year of publication: 1980 to 2009
Methodology: peer-reviewed article, health and safety agency report, conference proceedings, published in English
Activity: empirical study conducted in the laboratory or workplace, handling training or physical conditioning
Study objective: to evaluate the effectiveness of material handling training activities

Cochrane Group

Martimo et al., 2007 (n = 12)

Year of publication: 1981 to 2005
Methodology: randomized controlled trial or prospective cohort and case control study
Population: adult workers (16 to 70 years old, mixed), work activity: manual handling
Activity: changing behaviours, using handling aids
Study objective: to change manual handling behaviours; to measure back pain, disability related to back pain and absences due to illness

Verbeek et al., 2011 (n = 18)

Year of publication: 1981 to 2010
Methodology: randomized controlled trial or prospective cohort study with control group
Population: adult workers (16 to 70 years old, mixed), work activity: manual handling
Activity: changing behaviours, using handling aids
Study objective: to change manual handling behaviours; to measure back pain, disability related to back pain and absences due to illness

Other

Hogan et al., 2014 (n = 13)

Type of article: peer-reviewed article, published in English
Methodology: randomized controlled trial, quasi-experimental study or prospective cohort study with control or comparison group
Population: adult workers (16 to 70 years old, mixed), work activity: manual handling
Activity: targeted or comprehensive approach in the workplace, training in load handling or patient handling
Study objective: to prevent or reduce MSDs

Appendix 3 – Definitions of Analysis Variables

General variables related to the meta-analyses and the selection process for the studies selected

Variables	Classes	Clear mention or inference (M/I)	Descriptions
Key characteristics of the meta-analyses considered			
Date	*N/A*	M	Year of publication of the literature review
Number of articles	*N/A*	M	Number of articles selected for analysis by the literature reviews
Years covered	*N/A*	M	Years of inclusion of articles in literature reviews (e.g., from 1990 to 2014).
Objectives	*N/A*	M	Objectives of literature reviews
Results	*Positive* *Mixed* *Nil*	M	Results measured by literature reviews in the studies selected pour analysis (related to their evaluation objectives) *Positive: significant positive results or strong evidence of effectiveness* *Mixed: positive and nil results* *Nil: no significant positive result or no/little strong evidence*
Selection process based on articles referenced in the reviews			
Studies selected	*Yes* *No*	NA	Studies for which the analysis results are included in this review
Studies excluded	*Yes* *No*	NA	Studies for which the analysis results are not included in the review (e.g., literature reviews, groups of experts, duplicates, unclassified articles, etc.)
Environment	*Workplace (WP.)* *Laboratory (Lab.)* *Training Insitution (T.I.)*	M	Environment in which the training takes place *WP.: training occurs within the workplaces of various organizations (e.g., hospitals, storage and delivery companies)* *Lab.: training occurs in a "controlled environment" (e.g., biomechanical laboratory), making it possible to take complex measurements (e.g., EMG, kinematics), and the context is simplified (e.g., a single type of load, force platform that does not enable foot mobility)* *T.S.: training is provided in an educational institution (e.g., vocational school) for future professionals (e.g., nurses). Generally, manual handling is a module in their curriculum and training may include a work internship.*

Mention (M): element clearly mentioned; Clear inference (I): element not clearly mentioned, but which can be easily inferred from the other elements reported; N/A: not applicable

General variables of the studies selected

Variables	Classes	Clear mention or inference (M/I)	Descriptions
General characteristics of the studies			
Duration of the study	*≤ 1 month* *1 to 6 months* *6 to 12 months* *12 to 24 months* *> 24 months*	M	Total duration of the study (including the evaluations and the activity)
Objectives	*Reduced accidents/MSDs, pain* *Behaviour changes* *Improved physical capacity* *Improved knowledge* *Reduced health risk factors* *Changes* *Other*	M	Classes not mutually exclusive Objectives of the studies selected *Reduced accidents/MSDs/pain: evaluate the effectiveness of training in reducing/preventing workplace accidents (including costs, rates, severity, etc.), MSDs and pain* *Behaviour changes: evaluate the effectiveness of training on changing behaviour (e.g., adoption of the safe method, use of handling aids)* *Improved physical capacity: evaluate the effectiveness of training on improving physical capacities (e.g., strength, flexibility, cardiovascular endurance)* *Improved knowledge: evaluate the effectiveness of training on improving knowledge related to handling (e.g., injury mechanisms, risks, safe methods)* *Reduced health risk factors: evaluate the effectiveness of training in reducing risk factors related to health (e.g., lumbar loading, fatigue)* *Changes: evaluate the effectiveness of training on the implementation of changes (e.g., work environment, equipment, work organization)* *Other: evaluate the content of a training program by comparing it to recommendations from a recognized organization (e.g., compare the content to the recommendations in the manual handling guide published by the Chartered Society of Physiotherapy), evaluate the impact of training on participation in leisure activities, etc.*
Number of objectives	*1* *2* *≥ 3*	M	Number of objectives of the study (e.g., improvement of physical capacities, improved knowledge, changes, transformations)

Mention (M): element clearly mentioned; Clear inference (I): element not clearly mentioned, but which can be easily inferred from the other elements reported.

Variables related to the intervention

Variables	Classes	Clear mention or inference (M/I)	Descriptions
Characteristics of the organization			
Sector of activity	*Healthcare/ hospital* *Other sectors* *Multiple* *No specific sector*	M	Classes not mutually exclusive Sector of activity in which training takes place or from which the participants come or on which training is based *Other sectors: storage, construction, food sector, etc.* *Multiple: the participants come from more than one sector of activity* *No specific sector: does not target any particular work activity or sector of activity*
Number of sites	*One* *Several*	M or I	Number of independent worksites of the organization(s) (e.g., several warehouses, different care units/wings of a healthcare institution)
Organization size	*S* *M* *L* *VL*	M	Size of the organization where training took place *S: small organization (49 employees or less)* *M: medium-sized organization (50 to 499 employees)* *L: large organization (500 to 999 employees)* *VL: very large organization (1000 employees or more)*
Number of workplaces involved	*1* *2 to 9* *≥ 10*	M	Total number of organizations included in the study
Task(s) performed in the organization(s)	*Yes* *No*	M	Information about the tasks of employees of the organization(s) mentioned (e.g., tasks, types of loads, work schedule)
Types of loads handled in the scope of the work activity	*Patient* *Inert* *None*	M or I	Type of load handled in the scope of the work activity. Classes not mutually exclusive *Inert: box* *None: without relationship to any occupational activity (e.g., students)*

Mention (M): element clearly mentioned; Clear inference (I): element not clearly mentioned, but which can be easily inferred from the other elements reported.

Variables related to the intervention—continued

Variables	Classes	Clear mention or inference (M/I)	Descriptions
Characteristics of interventions			
Duration of the intervention	*≤ 1* *1 to 6* *6 to 12* *12 to 24* *> 24 months*	M	Total duration of the intervention (including the training and changes made)
Scope of the intervention	*Targeted* *Comprehensive*	M	*Targeted: training only* *Comprehensive: training follows a comprehensive risk prevention approach (e.g., addition of handling aids, changes to work organization, changes to the environment, evaluation of risks)*
Type of intervention	*Physical conditioning* *Training* *Change(s)*	M or I	Classes not mutually exclusive *Physical conditioning: improvement of physical capacities* *Training: knowledge transfer/change/behaviour change* *Change(s): targets changes in the organization (e.g., changes in OHS management, addition of handling aids, changes in the work environment)*
Phases reported	*Investigations* *Diagnosis* *Recommendations* *Changes* *Evaluation of risks* *Post-training follow-up* *Creation of a follow-up committee/work-ing group*	M	*Classes not mutually exclusive* *Investigations: work analysis to guide the intervention (investigations not used to build or adapt the training program)* *Diagnosis: presentation of a diagnosis following investigations* *Recommendations: presentation of recommendations* *Changes: changes in the work activity (e.g., layout, equipment, procedures)* *Evaluation of risks: analyses to identify/evaluate risks to health within the organization* *Post-training follow-up: evaluation of the training program by the trainer (e.g., results measured by questionnaire or observation after training ends, satisfaction of trainees, knowledge acquisition, etc.)* *Creation of a follow-up committee/working group: involvement in or striking of a committee/follow-up group or working group (support to trainers, follow-up of a change project, participation in the development of the training program)*
Information about the request	*Yes* *No*	M	Information about the initial request that led to the training program/intervention mentioned

Mention (M): element clearly mentioned; Clear inference (I): element not clearly mentioned, but which can be easily inferred from the other elements reported.

Variables related to the intervention—continued

Variables	Classes	Clear mention or inference (M/I)	Descriptions
Characteristics of participants			
Number of participants	*1 to 49* *50 to 99* *≥ 100*	M	Number of participants included in the study
Exposure of participants	*Same exposure* *Different exposure*	M	*Same exposure: the participants in the same group received the same training/activity* *Different exposure: the participants in the same group received different training/activities*
Gender	*Mixed* *Male* *Female*	M	Gender of participants
Age	*18 to 29* *30 to 39* *≥ 40 years*	M	Average age of participants
Information about minimum age	*Yes* *No*	M	Information about the minimum age of participants mentioned
Information about maximum age	*Yes* *No*	M	Information about the maximum age of participants mentioned
Seniority	*0 to 5* *5 to 10* *≥ 10 years*	M	Average seniority (in years) of participants in the organization
Information about minimum seniority	*Yes* *No*	M	Information about the minimum length of seniority of participants in the organization mentioned
Information about maximum seniority	*Yes* *No*	M	Information about the maximum length of seniority of participants in the organization mentioned
Health history	*Negative: no history* *Positive: with history of injury/MSD/pain*	M	*Negative: no health problem (pain, injury, MSD)* *Positive: has current or past health problems (pain, injury, MSD)*
Nature	*Pain* *Injuries/ MSDs/ Accidents* *Pain and injuries/ MSDs/Accidents*	M	Main problems among participants with a positive health history (e.g., injury, pain, MSD)
Site/body	*Back* *General*	M	Main injury site among participants with a positive health history *General: the entire body*
Percentage	*1 to 49%* *50 to 99%* *100%*	M	Percentage of participants with a positive health history

Mention (M): element clearly mentioned; Clear inference (I): element not clearly mentioned, but which can be easily inferred from the other elements reported.

Variables related to the intervention—continued

Variables	Classes	Clear mention or inference (M/I)	Descriptions
Characteristics of participants			
Profile	*Worker* *Student* *Manager* *Supervisor* *Other*	M	Classes not mutually exclusive Position/function within or outside the organization
Occupation	*Care staff* *Handler* *Housekeeper* *Other* *None*	M	Classes not mutually exclusive *Care staff: nurses, nursing assistants, attendants, physiotherapists, occupational therapists, etc.* *Handler: worker whose principal activity or one of his/her principal activities is handling inert loads (e.g., order picker, delivery person, etc.)* *Others: office staff, teachers, military personnel* *None: students with no specified field of study*
Homogeneity of occupations	*Homogeneous* *Heterogeneous*	M	*Homogeneous: the participants have the same occupation* *Heterogeneous: the participants have varied occupations*
Handling experience	*Yes* *No* *Mixed*	M or I	*Yes: the participants have previously performed handling tasks in the scope of a job* *No: no experience* *Mixed: some participants have experience*

Mention (M): element clearly mentioned; Clear inference (I): element not clearly mentioned, but which can be easily inferred from the other elements reported.

Variables related to the training programs

Variables	Classes	Clear mention or inference (M/I)	Descriptions
Characteristics of trainers			
Profile	*Occupational therapist/ Physiotherapist* *Ergonomist* *Kinesiologist* *Worker*	M	Classes not mutually exclusive Occupation/trainers' occupation (e.g., ergonomist, occupational therapist)
Origin	*Internally within the organization* *External to the organization*	M or I	Position of trainer(s) with respect to the organization *Internally within the organization: works within the organization (e.g., supervisor, worker)* *External to the organization: e.g., an expert or consultant mandated by the organization to carry out the activity (is not an employee or in upper management)*
Experience	*Yes* *No*	M	Information about the experience of the trainer mentioned
Number of trainers involved	*1* *≥ 2*	M	Number of trainers engaged in providing training
General characteristics of training programs			
Composition of groups	*Homogeneous* *Heterogeneous*	M or I	Composition of groups of participants during training sessions *Homogeneous: same hierarchical level (e.g., only handlers)* *Heterogeneous: participants with different hierarchical levels (e.g., workers and supervisors)*
Size of groups	*1* *2 to 5* *≥ 11*	M	Number of participants per group formed
Location	*Class* *Training room* *Simulated environment* *In the field*	M	Classes not mutually exclusive Location(s) where training takes place *Simulated environment: environment that simulates/imitates the actual work environment* *In the field: training conducted where the working activity normally takes place (e.g., care units, warehouse)* *Class: classroom*
Procedure	*Theory and practice* *Practice only* *Theory only* *Internship*	M	*Theory only: with theoretical segment only (no practical segment)* *Practice only: with motor engagement only (no theoretical segment)* *Theory and practice: includes both a theoretical and a practical segment* *Internship: educational training that includes a workplace internship*

Mention (M): element clearly mentioned; Clear inference (I): element not clearly mentioned, but which can be easily inferred from the other elements reported.

Variables related to the training programs—continued

Variables	Classes	Clear mention or inference (M/I)	Descriptions
General characteristics of training programs			
Session formats	*Several sessions* *A single session*	M	Training organization over time *A single session: each individual participated in a single session* *Several sessions: each individual participated in more than one session*
Number of sessions	*1* *2 to 4* *≥ 5*	M	Total number of sessions per training program (e.g., 1 session: each individual participated in a single session)
Total duration	*< 1 day* *1 day* *≥ 2 days*	M	Total duration of the training program in hours (theoretical and practical portions, 1 day = 8 hr., ½ day = 4 hr.)
Duration of the theoretical portion	*0* *< 1 day* *1 day* *≥ 2 days*	M	Total duration of the theoretical portion, i.e., without motor activation
Duration of the practical portion	*0* *< 1 day* *1 day* *≥ 2 days*	M	Total duration of the practical portion, i.e., with motor activation
Time span	*≤ 1 day* *2 to 30 days* *31 to 365 days* *1 year to 2 years* *≥ 2 years*	M	Number of days over which the training sessions took place
Time between sessions	*Yes* *No*	M	Information about the number of days separating the training sessions mentioned
Type of load used	*Patient* *Inert* *Other/None*	M or I	Classes not mutually exclusive Type of load used in the scope of training *Other/None: other loads were used (e.g., training equipment) or no load was used/discussed*
Number of teaching tools used	*1* *≥ 2*	M	Number of tools used for each training program (e.g., slideshow, workshops, etc.)
Teaching tools	*Workshops* *Visual component* *Participant's document* *Other*	M	Classes not mutually exclusive *Workshops: exercises and workshops (e.g., case studies, simulations)* *Visual component: visual support (e.g., slideshow, photos, videos)* *Participant's document: hard copy (e.g., training document)* *Other: educational tool that does not enter into the preceding categories (e.g., X-ray)*

Mention (M): element clearly mentioned; Clear inference (I): element not clearly mentioned, but which can be easily inferred from the other elements reported.

Variables related to the training programs—continued

Variables	Classes	Clear mention or inference (M/I)	Descriptions
Training program characteristics			
Recommended educational approach	*Top down/ knowledge transfer agent* *Mixed/transfers knowledge and encourages exchanges* *Bottom up/encourages exchanges*	I	Trainer's teaching approach / posture / attitude during the training program *Top down/knowledge transfer agent: the trainer acts as a knowledge transfer agent* *Mixed/transfers knowledge and encourages exchanges: combination of the two approaches* *Bottom up/encourages exchanges: exchanges and discussions based on the trainees' experiences, the trainer acts more as a facilitator*
Feedback	*Yes* *No*	M	The trainer does reinforcement activities with the participants (e.g., feedback, reminders about good practices, biofeedback)
Biofeedback	*Yes* *No*	M	Subcategory of "Feedback" Includes biofeedback, which is feedback based on the reactions of the nervous/physiological system that can be both auditory and visual
Interactions with participants	*Yes* *No*	M	Includes exchanges/discussions among the participants or with the trainer
Key principle guiding the training	*Safe method* *Principles of physical conditioning* *Ergonomics* *Risk evaluation* *Experts' strategies*	M or I	Classes not mutually exclusive Model/key principle on which training content and format is based *Safe method: often seen as the only method to use, e.g.: straight back and bent knees, keeping the load close to the body, being stable and using smooth and controlled movements* *Principles of physical conditioning: e.g., improving strength and endurance, stretching* *Ergonomics: e.g., considering the determinants of the work activity and the physical, cognitive and social aspects of the activity* *Risk assessment: identification/evaluation of risks in the organization or when tasks are performed as a preventive strategy (safe approach)* *Experts' strategies: based on strategies adopted by expert handlers*
Match between the loads handled at work and those used during training	*Yes* *No* *Without relationship to any occupational activity*	M or I	Loads used/discussed in the training program are the same as those mentioned in the actual work activity

Mention (M): element clearly mentioned; Clear inference (I): element not clearly mentioned, but which can be easily inferred from the other elements reported.

Variables related to the training programs—continued

Variables	Classes	Clear mention or inference (M/I)	Descriptions
Training characteristics			
Appropriateness	*Turnkey* *Adapted to environment* *Adapted to context*	M or I (M for "Adapted to context")	*Turnkey: not adapted to the environment, training is pre-designed or ready to be applied in any environment* *Adapted to environment: e.g., adapted to a hospital or manufacturing environment, without being adapted to the actual work context (e.g.: types of patients, production, demands)* Adapted to context: considers the specific context of the company/institution (e.g., type of patient, procedures, demands, difficulties experienced by the workers)
Includes preliminary analyses	*Yes* *No*	M	Subcategory of "Adapted to context" Preliminary analyses performed to design the training program (e.g., observations, interviews, analyses of the organization's documentation)
Includes illustrations	*Yes* *No*	M	Subcategory of "Adapted to context" Training includes images or videos taken from the actual work context
Knowledge transferred	*Knowledge* *Know-how* *Soft skills*	I	Classes not mutually exclusive *Knowledge: transfer/acquisition of knowledge/related to the activity or not* *Know-how: involves the transfer/acquisition of skills to be applied in the work activity* *Soft skills: knowledge built from the worker's characteristics and relationships within the community (e.g., principles/values to respect when performing a task, "paying attention")*
Type of training	*Knowledge transfer* *Physical conditioning* *Behaviour change* *Awareness raising*	M or I	Classes not mutually exclusive *Physical conditioning: uses or is solely based on carrying out an athletic training program or a handling task (e.g., stationary bicycle, repetition of a task outside of its context with the goal of increasing physical capacity)* *Behaviour change: change in motor behaviour, no judgments or decision-making by the trainee (simple application of operating methods without consideration of the context)* *Knowledge transfer: acquisition of know-how applicable in the context of work (strategies) involving judgment/decision making and practice of motor skills (e.g., practice of handling strategies in the field) or non-motor skills (e.g., use of the NIOSH equation)* *Awareness raising: information session without transfer of know-how* *Subcategories:* *"Back school": information session aimed at knowledge acquisition and behaviour change, often integrating physical conditioning exercises* *Biofeedback: based on actions/reactions of the nervous system (e.g., electrodes that perceive variations of muscle contractions when movements are performed and that send sound or visual signals to the participant)*

Mention (M): element clearly mentioned; Clear inference (I): element not clearly mentioned, but which can be easily inferred from the other elements reported.

Variables related to the training programs—continued

Variables	Classes	Clear mention or inference (M/I)	Descriptions
Characteristics of training programs			
Subjects covered			
Handling techniques	*Yes* *No*	M	Deals with general handling techniques/methods (subcategories if specified in the techniques/methods taught)
Safe handling technique or "good method"	*Yes* *No*	M	Integrates so-called safe handling methods (e.g., straight back and bent knees) or deals with the "good handling method" (with or without clarification of the principles presented)
Effects of handling on health	*Yes* *No*	M	Deals with the effects/consequences of handling on health and the theoretical elements related to it (e.g., biomechanics, anatomy, accidents/injuries/OHS statistics)
Biomechanics	*Yes* *No*	M	Subcategory of "Effects of handling on health" Biomechanical effects related to manual handling (e.g., loading on the lumbar region, shearing, compression, twisting)
Anatomy	*Yes* *No*	M	Subcategory of "Effects of handling on health" Theoretical concepts about the anatomy of the body related to the effects of manual handling (e.g., spinal column, vertebrae, nerves, discs)
Accidents/injuries/ statistics	*Yes* *No*	M	Subcategory of "Effects of handling on health" Statistics or examples of risks of accident/injuries/pain related to manual handling (e.g., herniated discs, lumbar sprains, low back pain, frequent accidents in the work environment)
Determinants of the work activity	*Yes* *No*	M	Concepts of the determinants of the work activity (e.g., environmental characteristics, loads, tools used, procedures)
Handling aids	*Yes* *No*	M	Subcategory of "Work activity determinants" *Handling aids covered during the training program (e.g., carts, lift system)*
Environmental characteristics	*Yes* *No*	M	Subcategory of "Work activity determinants" *Characteristics of the environment covered during the training program (e.g., space available, deposit height)*
Load characteristics	*Yes* *No*	M	Subcategory of "Work activity determinants" *Characteristics of loads/patients handled discussed during the training program (e.g., weight, fragility, independence of the patient)*
Physical conditioning principles	*Yes* *No*	M	Elements related to improving muscular or cardiovascular fitness (e.g., flexibility, cardiovascular or strength training)
Pain/stress management	*Yes* *No*	M	Strategies aimed at managing/decreasing/preventing pain/injuries, discomfort, psychological stress (e.g., stretching, breathing techniques)

Mention (M): element clearly mentioned; Clear inference (I): element not clearly mentioned, but which can be easily inferred from the other elements reported.

Variables related to the training programs—continued

Variables	Classes	Clear mention or inference (M/I)[1]	Descriptions
Characteristics of training programs			
Subjects covered			
Risk factors for pain/injuries/accidents	*Yes* *No*	M	Risk factors and their identification/elimination/reduction are covered, using a safe approach (e.g., identification method, identification tools, load weights, deposit height)
Prevention at work	*Yes* *No*	M	Prevention in OHS (e.g., methods taken to avoid/reduce the incidence/seriousness of accidents/MSD/pain)
Work activity	*Yes* *No*	M	Work activity of the participants for their physical, social and cognitive dimensions (e.g., tasks, collective aspects, difficulties experienced)
Organization of work	*Yes* *No*	M	Subcategory of "Work activity" Elements related to the principles of organization of work (e.g., avoidance of rehandling, distribution of efforts, planning)
Planning	*Yes* *No*	M	Subcategory of "Organization of work" Planning of tasks and operations related to handling
Communication with coworkers	*Yes* *No*	M	Subcategory of "Work activity" *Concept of communication/transmission of information among workers in the scope of their jobs (e.g., importance of communication during handling by two people, exchanging strategies, etc.)*
Sharing expertise	*Yes* *No*	M	Subcategory of "Work activity" Elements related to the transfer of knowledge and skills to coworkers (e.g., transmission strategies, concepts to be transmitted, etc.)
Difficulties experienced	*Yes* *No*	M	Subcategory of "Work activity" Integrates/addresses the real difficulties experienced by trainees in the scope of their jobs (e.g., difficulties related to handling of a certain type of patient, difficulties related to use of a cart, etc.)
OHS management/regulations/laws	*Yes* *No*	M	Addresses management/administration concepts related to OHS (e.g., use of certain approaches/tools such as PDCA[2]).
Changes at work	*Yes* *No*	M	Integrates the concept of changes in the work activity (e.g., retrofitting, purchase of equipment, etc.), their possible impact, the role of workers, etc.
Individual aspects	*Yes* *No*	M	Addresses concepts related to the individual/worker, e.g., his/her physical characteristics (strength, flexibility), responsibility with regard to OHS (using the techniques that were taught).

[1.] Mention (M): element clearly mentioned; Clear inference (I): element not clearly mentioned, but which can be easily inferred from the other elements reported; [2.] PDCA: Plan Do Check Act

Variables related to the evaluation of the training program by the authors of the studies selected

Variables	Classes	Clear mention or inference (M/I)	Descriptions
Characteristics of the evaluation			
Evaluation model	*Experimental/ quasi-experimental* *Descriptive/by observation*	M or I	*Experimental/quasi-experimental: validation and implementation of the training program/activity by the research team; use of an appropriate sampling method (e.g., use of control groups, random selection and assignment of participants to groups); evaluation tools validated, etc.* *Descriptive/by observation: evaluation of a training program/activity not validated or set up/implemented by the research team. The team observes the effects of an activity previously carried out or in progress with a given population. May include a control group, cross-sectional or longitudinal data collection via questionnaire.*
Number of groups	*1* *2* *≥ 3*	M	Number of groups of participants included in the study
Control group	*Yes* *No*	M	Setting up of a control group
Control group exposed	*Yes* *No*	M	Subcategory of "Control group" Group considered as a "control" by the authors and which had received training.
Number of protocols	*1* *2* *≥ 3*	M	Number of training programs with different content and formats evaluated by the study (that were not used as a control group)
Collection period	*Pre-* *Post-* *Follow-up* *Other*	M	Classes not mutually exclusive Moments/evaluation/result collection period *Pre-: data collection prior to the training program/activity* *Post-: data collection directly after the training program/activity* *Follow-up: latency period between the end of the training program/activity and data collection (different than Post-)* *Other: data collection other than post-/follow-up/pre- (e.g., ongoing, mid-training/activity, etc.)*
Measurement design	*Pre-/post-* *Pre-follow-up* *Post-* *Follow-up*	M	Classes not mutually exclusive *Pre-post-: post-training/activity data compared to pre-training/activity data* *Pre-follow-up: data collected during follow-up compared to pre-training/activity data* *Post-: post-training/activity data as the only evaluation (e.g., data that can be compared with a control group)* *Follow-up: use of data collected during the follow-up of the training program/activity as the only evaluation (e.g., can be compared to a control group)*
Number of distinct evaluations	*1* *2* *3*	M	Total number of post-training evaluations and follow-ups that could have different latency periods

Mention (M): element clearly mentioned; Clear inference (I): element not clearly mentioned, but which can be easily inferred from the other elements reported.

Variables related to the evaluation of the training program by the authors of the studies selected—continued

Variables	Classes	Clear mention or inference (M/I)	Descriptions
Characteristics of the evaluation			
Latency period between pre- and post-measurements	*0 to 1* *1 to 6* *6 to 12* *≥ 12 months*	M	Latency period between the pre-training program/activity and post-training program/activity (in months)
Latency period between pre- and follow-up	*0 to 1* *1 to 6* *6 to 12* *12 to 24* *≥ 24 months*	M	Classes not mutually exclusive Latency period between the pre-training program/activity and the follow-up after the training program/activity (in months)
Latency period between the end of training and the post- and/or follow up measurement	*0* *0 to 1* *1 to 6* *6 to 12* *≥ 12 months*	M	Classes not mutually exclusive Latency period separating the end of training and the post-training program/activity evaluations and the follow-ups (month)
Length of follow-up	*≤ 1 year* *> 1 ear*	M	Duration over which data were compiled continuously (different from pre-follow-up, e.g., accident statistics)
Number of evaluation tools used	*1* *2* *≥ 3*	M	Number of tools used by the authors to evaluate the training
Evaluation tools used	*Questionnaire* *Observations* *Physical capacity tests* *OHS statistics* *Interviews* *Other*	M	Classes not mutually exclusive *Questionnaires:* document that includes a series of questions to be answered by the participants *Observations:* observations of the population being evaluated Subcategories: In the actual context: observation in the actual work environment and its context (e.g., care unit, warehouse, etc.) In a simulated environment: observation in a simulated environment or not in the actual context *Physical capacity tests:* evaluation of the VO_2max, physical endurance, maximum strength, etc. *OHS statistics:* e.g., the organization's accident book, government records, etc. *Interviews:* individual or collective meetings about the work methods used, the restrictions experienced by the employees etc. *Other:* audit, NIOSH lifting equation and organization documents (e.g., management system)
Review by the authors	*Positive assessment* *Mixed assessment* *No assessment*	M	Review by the authors of the effectiveness of the training program(s) evaluated in their study (self-reported) *Positive assessment: the review found that training was effective* *Mixed assessment: the review found that training was effective according to some indicators, but had no effect over one or several indicators* *No assessment: the review did not find the training program effective*

Mention (M): element clearly mentioned; Clear inference (I): element not clearly mentioned, but which can be easily inferred from the other elements reported.

Variables related to the evaluation of the training program by the authors of the studies selected—continued

Variables	Classes	Clear mention or inference (M/I)	Descriptions
Characteristics of the evaluation			
Indicators used to evaluate the training program: general indicators			
Changes measured among the participants	*Significant positive results* *Mixed results* *Nil results* *Negative results* *N/A*	M	E.g., behaviour change, enhanced knowledge, improved physical capacities, etc. *Significant positive results: the indicator(s) presented positive and significant results* *Mixed results: some indicators with positive results and others with nil or negative results* *Nil results: no positive and significant effect measured* *Negative results: detrimental/aggravating effects measured* *N/A: not evaluated by the study*
Consequences related to handling	*Significant positive results* *Mixed results* *Nil results* *Negative results* *N/A*	M	Measured results related to the possible consequences of the work activity on the individual/organization (e.g., reduction in injuries, reduction in accident costs, etc.) *Significant positive results: the indicator(s) presented positive and significant results* *Mixed results: some indicators with positive results and others with nil or negative results* *Nil results: no positive and significant effect measured* *Negative results: detrimental/aggravating effects measured* *N/A: not evaluated by the study*
Reduction of health risk factors	*Significant positive results* *Mixed results* *Nil results* *Negative results* *N/A*	M	Measured results related to the reduction of risk factors in the work activity (e.g., risks of MSDs, fatigue, lumbar loading, etc.) *Significant positive results: the indicator(s) presented positive and significant results* *Mixed results: some indicators with positive results and others with nil or negative results* *Nil results: no positive and significant effect measured* *Negative results: detrimental/aggravating effects measured* *N/A: not evaluated by the study*
Training quality	*High* *Average* *Poor* *N/A*	M	Measured results related to the quality of training (e.g., content, teaching methods, paradigm, etc.) *High: quality of training assessed overall as good/very good* *Average: quality of training assessed overall as average (some elements are of high quality while others are of low quality)* *Poor: quality of training assessed overall as bad/inadequate*
Cost-benefits ratio	*Significant positive results* *Mixed results* *Nil results* *Negative results* *N/A*	M	Measured results related to the comparison between the financial costs of training programs/activities and their benefits (e.g., reduction in accident rates or reduction in accident costs) *Significant positive results: the indicator(s) presented positive and significant results* *Mixed results: some indicators with positive results and others with nil or negative results* *Nil results: no positive and significant effect measured* *Negative results: detrimental/aggravating effects measured* *N/A: not evaluated by the study*

Mention (M): element clearly mentioned; Clear inference (I): element not clearly mentioned, but which can be easily inferred from the other elements reported.

Variables related to the evaluation of the training program by the authors of the studies selected—continued

Variables	Classes	Clear mention or inference (M/I)	Descriptions
Characteristics of the evaluation			
Indicators used to evaluate the training program: subcategories of "Changes measured among the participants"			
Behaviour changes	*Significant positive results* *Mixed results* *Nil results* *Negative results* *N/A*	M or I	Measured results related to participants' behavioural changes. Requires no know-how in terms of the context of the work activity (no decision-making based on the context, e.g., adoption of the safe method, use of handling aids, postural changes, etc.) *Significant positive results; Mixed results; Nil results; Negative results; N/A: same as "Changes measured in the participant"*
Improvement of physical capacities	*Significant positive results* *Mixed results* *Nil results* *Negative results* *N/A*	M	Measured results related to the improvement of participants' physical capacities (e.g., increase in cardiovascular capacities, increased strength, etc.) *Significant positive results; Mixed results; Nil results; Negative results; N/A: same as "Changes measured in the participant"*
Satisfaction level reported by the participants	*Significant positive results* *Mixed results* *Nil results* *Negative results* *N/A*	M	Measured results related to participants' satisfaction with the training *Significant positive results; Mixed results; Nil results; Negative results; N/A: same as "Changes measured in the participant"*
Improvement of skills	*Significant positive results* *Mixed results* *Nil results* *Negative results* *N/A*	M or I	Measured results related to the improvement of participants' skills: know-how and judgement (e.g., ability to assess risks, choice of operating methods according to the context, recommendations made, etc.) *Significant positive results; Mixed results; Nil results; Negative results; N/A: same as "Changes measured in the participant"*
Improvement of knowledge	*Significant positive results* *Mixed results* *Nil results* *Negative results* *N/A*	M	Measured results related to the improvement of participants' knowledge. No know-how (e.g., injury risk factors, identification of the safe method, etc.) *Significant positive results; Mixed results; Nil results; Negative results; N/A: same as "Changes measured in the participant"*
Others	*Significant positive results* *Mixed results* *Nil results* *Negative results* *N/A*	M	Measured results related to attitudinal change in terms of risk, sense of individual responsibility, participation in physical leisure activities and perception of work *Significant positive results; Mixed results; Nil results; Negative results; N/A: same as "Changes measured in the participant"*

Mention (M): element clearly mentioned; Clear inference (I): element not clearly mentioned, but which can be easily inferred from the other elements reported.

Variables related to the evaluation of the training program by the authors of the studies selected—continued

Variables	Classes	Clear mention or inference (M/I)	Descriptions
Characteristics of the evaluation			
Indicators used to evaluate the training program: subcategories of "Consequences related to handling"			
Reduction in pain	*Significant positive results* *Mixed results* *Nil results* *Negative results* *N/A*	M	Measured results related to the pain experienced by participants (e.g., low back pain, arm pain, back pain, etc.) *Significant positive results: the indicator(s) presented positive and significant results* *Mixed results: some indicators with positive results and others with nil or negative results* *Nil results: no positive and significant effect measured* *Negative results: detrimental/aggravating effects measured* *N/A: not evaluated by the study*
Reduction in injuries	*Significant positive results* *Mixed results* *Nil results* *Negative results* *N/A*	M	Measured results related to workplace / occupational injuries (e.g., MSDs and accidental events) *Significant positive results: the indicator(s) presented positive and significant results* *Mixed results: some indicators with positive results and others with nil or negative results* *Nil results: no positive and significant effect measured* *Negative results: detrimental/aggravating effects measured* *N/A: not evaluated by the study*
Reduction in accident costs	*Significant positive results* *Mixed results* *Nil results* *Negative results* *N/A*	M	Measured results related to the costs of workplace accidents (claims, direct and indirect costs, etc.) *Significant positive results: the indicator(s) presented positive and significant results* *Mixed results: some indicators with positive results and others with nil or negative results* *Nil results: no positive and significant effect measured* *Negative results: detrimental/aggravating effects measured* *N/A: not evaluated by the study*

Mention (M): element clearly mentioned; Clear inference (I): element not clearly mentioned, but which can be easily inferred from the other elements reported.

Variables related to the evaluation of the training program by the authors of the studies selected—continued

Variables	Classes	Clear mention or inference (M/I)	Descriptions
Characteristics of the evaluation			
Indicators based on Kirkpatrick's evaluation model (1994)			
Number of levels used	*1* *2* *3* *4*	I	Number of levels used by the study, without taking into account the sequence between levels
Level 1: satisfaction	*Significant positive results* *Mixed results* *Nil results* *Negative results*	I	Measured results related to trainees' satisfaction with the training program/activity (may include an assessment of reduction in discomfort, fatigue, pain, if assessed by questionnaire) *Significant positive results: effectiveness of training on this category of indicators* *Mixed results: training is effective for some indicators and not effective for others* *Nil results: no effect of training on this category of indicators* *Negative results: training has negative/detrimental effects on this category of variables*
Level 2: learning	*Significant positive results* *Mixed results* *Nil results* *Negative results*	I	Measured results related to trainees' learning (knowledge acquisition, behaviour change, skills improvement, improvement of physical abilities and attitudinal change) outside of the actual work context *Significant positive results: effectiveness of training on this category of indicators* *Mixed results: training is effective for some indicators and not effective for others* *Nil results: no effect of training on this category of indicators* *Negative results: training has negative/detrimental effects on this category of variables*
Level 3: transfer	*Significant positive results* *Mixed results* *Nil results* *Negative results*	I	Measured results related to the transfer of learning in the actual workplace (real context) (e.g., use of assistive devices, choice of handling techniques, etc.), can include behavioural changes and skills improvement *Significant positive results: effectiveness of training on this category of indicators* *Mixed results: training is effective for some indicators and not effective for others* *Nil results: no effect of training on this category of indicators* *Negative results: training has negative/detrimental effects on this category of variables*

Mention (M): element clearly mentioned; Clear inference (I): element not clearly mentioned, but which can be easily inferred from the other elements reported.

Variables related to the evaluation of the training program by the authors of the studies selected—continued

Variables	Classes	Clear mention or inference (M/I)	Descriptions
Characteristics of the evaluation			
Indicators based on Kirkpatrick's evaluation model—continued (1994)			
Level 4: Results	*Significant positive results* *Mixed results* *Nil results* *Negative results*	I	Measured results related to the "results" obtained from the training program/activity within the company/institution (e.g., improved production, accident reduction, reduction in accident-related costs, reduction in staff turnover, etc.) *Significant positive results: effectiveness of training on this category of indicators* *Mixed results: training is effective for some indicators and not effective for others* *Nil results: no effect of training on this category of indicators* *Negative results: training has negative/detrimental effects on this category of variables*
Respect of the sequence	*Yes* *No*	I	Studies that respected Kirkpatrick's (1994) sequence of levels
Number of levels reached	*1st level* *1st and 2nd level* *1st to the 3rd level* *1st to the 4th level*	I	*1st level: only the 1st level used/reached* *1st and 2nd levels: levels 1 and 2 were reached/used* *1st to the 3rd level: levels 1, 2 and 3 were reached/used* *1st to the 4th level: levels 1, 2, 3 and 4 were reached/used*

Mention (M): element clearly mentioned; Clear inference (I): element not clearly mentioned, but which can be easily inferred from the other elements reported.

Variables related to the evaluation of the training program by the authors of the studies selected—continued

Variables	Classes	Clear mention or inference (M/I)	Descriptions
Characteristics of the evaluation according to literature reviews			
Classification of the methodological quality of articles by the reviews	*High quality* *Other quality* *Unassigned* *Contradictory rating*	M	*High quality: the reviews characterized the studies as having high methodological quality* *Other quality: the reviews characterized the studies as having a good, average or poor methodological quality* *Unassigned: no evaluation of the methodological quality was carried out* *Contradictory rating: at least two reviews obtained different results (e.g., high vs. insufficient quality) in the evaluation of the methodology of a single study* *HSE Group and Hogan et al. (2014):* *QR 0 – 69%: Other quality* *QR 0 – 49%: Weak quality* *QR 50 – 59%: Average quality* *QR 60 – 69%: Good quality* *QR ≥ 70%: High quality* *Cochrane Group:* *High quality: responds to at least 7 criteria* *Poor quality: responds to 6 criteria or less*

Mention (M): element clearly mentioned; Clear inference (I): element not clearly mentioned, but which can be easily inferred from the other elements reported.

Appendix 4 – Additional Variables Not Included in the Results Tables

General characteristics of the studies selected

Studies selected	WP. (n = 51)	Lab. (n = 17)	T.I. (n = 9)	Total (N = 77)
Number of objectives				
1	34 (67%)	15 (88%)	3 (33%)	52 (68%)
2	11 (22%)	2 (12%)	4 (44%)	17 (22%)
≥ 3	6 (12%)	-	2 (22%)	8 (10%)
Duration of the study				
≤ 1 month	5 (10%)	9 (53%)	1 (11%)	15 (19%)
1 to 6 months	10 (20%)	6 (35%)	-	16 (21%)
6 to 12 months	9 (18%)	-	-	9 (12%)
12 to 24 months	12 (24%)	-	-	12 (16%)
> 24 months	7 (14%)	-	3 (33%)	10 (13%)
Characteristics of participants				
Number of participants				
1 to 49	22 (43%)	15 (88%)	2 (22%)	39 (51%)
50 to 99	8 (16%)	1 (6%)	3 (33%)	12 (16%)
≥ 100	17 (33%)	-	3 (33%)	20 (26%)
Gender				
Mixed	22 (43%)	5 (29%)	4 (44%)	31 (40%)
Male	3 (6%)	10 (59%)	-	13 (17%)
Female	6 (12%)	2 (12%)	2 (22%)	10 (13%)
Age				
18 to 29	4 (8%)	9 (53%)	2 (22%)	15 (19%)
30 to 39	13 (25%)	1 (6%)	1 (11%)	15 (19%)
≥ 40 years	6 (12%)	-	-	6 (8%)
Health history				
Negative: without history	7 (14%)	12 (71%)	-	19 (25%)
Positive: with injury/MSD/pain history	20 (39%)	1 (6%)	4 (44%)	25 (32%)
Occupation[1]				
Care staff	29 (57%)	1 (6%)	9 (100%)	39 (51%)
Material handler	16 (31%)	3 (18%)	-	19 (25%)
Housekeeper	4 (8%)	-	-	4 (5%)
Other[2]	11 (22%)	2 (12%)	-	13 (17%)
None[3]	-	8 (47%)	-	8 (10%)
Homogeneity of occupations				
Homogeneous	24 (47%)	12 (71%)	8 (89%)	44 (57%)
Heterogeneous	24 (47%)	1 (6%)	1 (11%)	26 (34%)
Material handling experience				
Yes	29 (57%)	3 (18%)	3 (33%)	35 (45%)
No	-	8 (47%)	-	8 (10%)
Mixed	3 (6%)	-	2 (22%)	5 (6%)

[1] *Not mutually exclusive subcategories.*
[2] *Other: may include office staff, teachers and military personnel.*
[3] *None: students without any specified field of study.*

General characteristics of the studies selected—continued

Studies selected	WP. (n = 51)	Lab. (n = 17)	T.I. (n = 9)	Total (N = 77)
Characteristics of the intervention				
Scope of the intervention				
Targeted	34 (67%)	17 (100%)	9 (100%)	60 (78%)
Comprehensive	17 (33%)	-	-	17 (22%)
Training characteristics				
Key principle guiding the training program[1,3]				
Safe method	46 (90%)	7 (41%)	8 (89%)	61 (79%)
Principles of physical conditioning	11 (22%)	9 (53%)	-	20 (26%)
Experts' strategies	-	1 (6%)	-	1 (1%)
Risk evaluation	9 (18%)	1 (6%)	1 (11%)	11 (14%)
Correspondence to the work activity[3]				
Turnkey	14 (27%)	10 (59%)	-	24 (31%)
Adapted to environment	16 (31%)	8 (47%)	9 (100%)	33 (43%)
Adapted to context	12 (24%)	-	-	12 (16%)
Type of training[1]				
Knowledge transfer	19 (37%)	1 (6%)	7 (78%)	27 (35%)
Physical conditioning	8 (16%)	10 (59%)	-	18 (23%)
Behaviour change	15 (29%)	6 (35%)	-	21 (27%)
Awareness-raising	7 (14%)	5 (29%)	1 (11%)	13 (17%)
Composition of groups [2]				
Homogenous	36 (71%)	15 (88%)	9 (100%)	60 (78%)
Heterogeneous	7 (14%)	-	-	7 (9%)
Group size[3]				
1	5 (10%)	3 (18%)	-	8 (10%)
2 to 5	2 (4%)	1 (6%)	1 (11%)	4 (5%)
≥ 11	6 (12%)	-	1 (11%)	7 (9%)
Location[3]				
Training room/laboratory only	3 (6%)	17 (100%)	-	20 (26%)
In class only	18 (35%)	-	1 (11%)	19 (25%)
+ in a training room/laboratory	4 (8%)	-	-	4 (5%)
+ in a simulated environment	3 (6%)	-	1 (11%)	4 (5%)
+ in a simulated environment and training room	-	-	1 (11%)	1 (1%)
+ in the field	13 (25%)	-	3 (33%)	16 (21%)
In the field only	2 (4%)	-	-	2 (3%)
Session formats[3]				
A single session	12 (24%)	5 (29%)	2 (22%)	19 (25%)
Several sessions	30 (59%)	12 (71%)	3 (33%)	45 (58%)
Number of sessions				
2 to 4	18 (35%)	2 (12%)	2 (22%)	22 (29%)
≥ 5	12 (24%)	10 (59%)	-	22 (29%)
Procedure[3]				
Theory and practice	28 (55%)	4 (24%)	7 (78%)	39 (51%)
Practice only	7 (14%)	12 (71%)	1 (11%)	20 (26%)
Theory only	9 (18%)	1 (6%)	1 (11%)	11 (14%)

[1] *Not mutually exclusive subcategories.*
[2] *Hierarchical levels.*

General characteristics of the studies selected—continued

Studies selected	WP. (n = 51)	Lab. (n = 17)	T.I. (n = 9)	Total (N = 77)
Trainers' characteristics				
Trainer's profile[1]				
Occupational therapist/Physiotherapist	15 (29%)	-	-	15 (19%)
Ergonomist	6 (12%)	1 (6%)	-	7 (9%)
Kinesiologist	2 (4%)	1 (6%)	-	3 (4%)
Worker	2 (4%)	-	-	2 (3%)
Number of trainers involved				
1	14 (27%)	3 (18%)	-	17 (22%)
≥ 2	10 (20%)	-	-	10 (13%)
Training characteristics				
Duration of training				
Total				
< 1 day	13 (25%)	5 (29%)	2 (22%)	20 (26%)
1 day	8 (16%)	3 (18%)	1 (11%)	12 (16%)
≥ 2 days	10 (20%)	4 (24%)	2 (22%)	16 (21%)
Theory[2]				
0	5 (10%)	11 (65%)	-	16 (21%)
< 1 day	8 (16%)	1 (6%)	3 (33%)	12 (16%)
1 day	3 (6%)	-	-	3 (4%)
Practice[2]				
0	10 (20%)	2 (12%)	-	12 (16%)
< 1 day	4 (8%)	5 (29%)	3 (33%)	12 (16%)
1 day	3 (6%)	3 (18%)	-	6 (8%)
≥ 2 days	5 (10%)	3 (18%)	-	8 (10%)
Teaching tools used[1]				
Workshop	14 (27%)	7 (41%)	5 (56%)	26 (34%)
Visual component	14 (27%)	4 (24%)	1 (11%)	19 (25%)
Participants' document	11 (22%)	1 (6%)	2 (22%)	14 (18%)
Type of load used[1]				
Inert	20 (39%)	13 (76%)	-	33 (43%)
Patient	20 (39%)	1 (6%)	7 (78%)	28 (36%)
None or other	12 (24%)	2 (12%)	-	14 (18%)
Trainer's approach				
Top-down / knowledge transfer agent	32 (63%)	17 (100%)	4 (44%)	53 (69%)
Mixed / knowledge transfer agent and encouragement of exchanges between and among the participants	8 (16%)	-	1 (11%)	9 (12%)
Feedback	16 (31%)	2 (12%)	2 (22%)	20 (26%)
Biofeedback	1 (2%)	2 (12%)	-	3 (4%)
Interactions with participants	10 (20%)	-	1 (11%)	11 (14%)
Post-training follow-up	2 (4%)	1 (6%)	1 (11%)	4 (5%)
Knowledge transfer[1]				
Knowledge	43 (84%)	7 (41%)	7 (78%)	57 (74%)
Know-how	23 (45%)	2 (12%)	7 (78%)	32 (42%)

[1] *Not mutually exclusive subcategories.*

General characteristics of the studies selected—continued

Studies selected	WP. (n = 51)	Lab. (n = 17)	T.I. (n = 9)	Total (N = 77)
Training characteristics				
Subjects covered[1]				
Manual handling techniques	41 (80%)	6 (35%)	8 (89%)	55 (71%)
Safe handling techniques or "good method"	39 (76%)	6 (35%)	8 (89%)	53 (69%)
Effects of handling on health[1]	25 (49%)	3 (18%)	4 (44%)	32 (42%)
Biomechanics	19 (37%)	3 (18%)	3 (33%)	25 (32%)
Anatomy	11 (22%)	-	3 (33%)	14 (18%)
Accidents/injuries/statistics	10 (20%)	-	1 (11%)	11 (14%)
Determinants of the work activity[1]	25 (49%)	1 (6%)	5 (56%)	31 (40%)
Handling aids	21 (41%)	-	4 (44%)	25 (32%)
Environmental characteristics	5 (10%)	1 (6%)	1 (11%)	7 (9%)
Load characteristics	6 (12%)	-	1 (11%)	7 (9%)
Work activity[1]	14 (27%)	-	2 (22%)	16 (21%)
Organization of work	6 (12%)	-	1 (11%)	7 (9%)
Communication	3 (6%)	-	2 (22%)	5 (6%)
Transfer of knowledge to co-workers	2 (4%)	-	2 (22%)	4 (5%)
Difficulties experienced at work	2 (4%)	-	-	2 (3%)
Principles of physical conditioning	17 (33%)	10 (59%)	1 (11%)	28 (36%)
Pain/stress management	11 (22%)	1 (6%)	1 (11%)	13 (17%)
Pain/injury/accident risk factors	13 (25%)	1 (6%)	1 (11%)	15 (19%)
Injury prevention in the workplace	7 (14%)	-	1 (11%)	8 (10%)
OHS management/standards/laws	4 (8%)	-	3 (33%)	7 (9%)
Changes in the workplace[2]	4 (8%)	1 (6%)	-	5 (6%)
Individual aspects[3]	3 (6%)	-	-	3 (4%)

[1] *Not mutually exclusive subcategories.*
[2] *Possible impacts, participants' roles, etc.*
[3] *Health, physical capacities, role in relation to OHS, etc.*

Methodology of the studies selected

Studies selected	WP. (n = 51)	Lab. (n = 17)	T.I. (n = 9)	Total (N = 77)
Procedure used				
Evaluation model				
Experimental/quasi-experimental	39 (76%)	17 (100%)	3 (33%)	59 (77%)
Descriptive/by observation	12 (24%)	-	6 (67%)	18 (23%)
Control group	32 (63%)	9 (53%)	2 (22%)	43 (56%)
Control group exposed[1]	8 (16%)	1 (6%)	2 (22%)	11 (14%)
Measurement design[2]				
Pre-/post-	11 (22%)	11 (65%)	2 (22%)	24 (31%)
Pre-follow-up	26 (51%)	4 (24%)	4 (44%)	34 (44%)
Post-	1 (2%)	2 (12%)	1 (11%)	4 (5%)
Follow-up	22 (43%)	1 (6%)	5 (56%)	28 (36%)
Number of separate evaluations				
1	39 (76%)	15 (88%)	7 (78%)	61 (79%)
2	11 (22%)	2 (12%)	1 (11%)	14 (18%)
3	1 (2%)	-	1 (11%)	2 (3%)
Number of evaluation tools used				
1	15 (29%)	13 (76%)	7 (78%)	35 (45%)
2	21 (41%)	2 (12%)	1 (11%)	24 (31%)
≥ 3	15 (29%)	2 (12%)	1 (11%)	18 (23%)
Types of tools used[2]				
Questionnaires	32 (63%)	2 (12%)	8 (89%)	42 (55%)
Observations	22 (43%)	5 (29%)	3 (33%)	30 (39%)
Physical capacity tests	12 (24%)	11 (65%)	-	23 (30%)
OHS statistics	18 (35%)	-	-	18 (23%)
Interviews	8 (16%)	-	-	8 (10%)
Other	4 (8%)	1 (6%)	-	5 (6%)
Latency period				
Latency period between the pre- and post- and pre- and follow-up phases[2]				
0 to 1 month	5 (10%)	7 (41%)	-	12 (16%)
1 to 6 months	9 (18%)	6 (35%)	-	15 (19%)
6 to 12 months	7 (14%)	1 (6%)	-	8 (10%)
12 to 24 months	13 (25%)	-	1 (11%)	14 (18%)
≥ 24 months	7 (14%)	-	3 (33%)	10 (13%)
Latency period between the measurement taken at the end of training and the post- and/or follow-up measurement[2]				
0	11 (22%)	12 (71%)	2 (22%)	25 (32%)
0 to 1 month	4 (8%)	2 (12%)	2 (22%)	8 (10%)
1 to 6 months	5 (10%)	1 (6%)	1 (11%)	7 (9%)
6 to 12 months	6 (12%)	-	-	6 (8%)
≥ 12 months	9 (18%)	-	2 (22%)	11 (14%)
Undefined	21 (41%)	4 (23%)	5 (56%)	30 (39%)

[1]*Group of participants reported as being control subjects by the authors, but who received training.*
[2] *Not mutually exclusive subcategories.*

Results reported by the authors of the studies selected in terms of training effectiveness

Studies selected	WP. (n = 51)	Lab. (n = 17)	T.I. (n = 9)	Total (N = 77)
Results specific to each objective (formulated by the authors)[1]				
Reduction in accidents/MSDs/pain	**33 (65%)**	**1 (6%)**	**4 (44%)**	**38 (49%)**
Positive assessment	13 (25%)	-	1 (11%)	14 (18%)
Mixed assessment	-	-	1 (11%)	1 (1%)
Nil assessment	20 (39%)	1 (6%)	2 (22%)	23 (30%)
Behaviour changes	**14 (27%)**	**2 (12%)**	**5 (56%)**	**21 (27%)**
Positive assessment	10 (20%)	2 (12%)	3 (33%)	15 (19%)
Nil assessment	4 (8%)	-	2 (22%)	6 (8%)
Improvement of physical capacities	**6 (12%)**	**9 (53%)**	**-**	**15 (19%)**
Positive assessment	6 (12%)	9 (53%)	-	15 (19%)
Improvement of knowledge	**8 (16%)**	**1 (6%)**	**5 (56%)**	**14 (18%)**
Positive assessment	5 (10%)	1 (6%)	3 (33%)	9 (12%)
Mixed assessment	1 (2%)	-	-	1 (1%)
Nil assessment	2 (4%)	-	2 (22%)	4 (5%)
Reduction in health risk factors	**6 (12%)**	**4 (24%)**	**-**	**10 (13%)**
Positive assessment	4 (8%)	2 (12%)	-	6 (8%)
Mixed assessment	1 (2%)	-	-	1 (1%)
Nil assessment	1 (2%)	2 (12%)	-	3 (4%)
Other	**7 (14%)**	**2 (12%)**	**3 (33%)**	**12 (16%)**
Positive assessment	3 (6%)	-	1 (11%)	4 (5%)
Mixed assessment	1 (2%)	1 (6%)	-	2 (3%)
Nil assessment	3 (6%)	1 (6%)	2 (22%)	6 (8%)

[1] *Objectives not mutually exclusive.*

Effectiveness of workplace training (n = 48) according to whether or not the quality criteria used in this study are taken into account in training

Results	Criterion 1		Criterion 2		Criterion 3		Criterion 4	
	Yes (n = 12)	**No** (N = 36)	**Yes** (n = 32)	**No** (N = 16)	**Yes** (n = 18)	**No** (N = 30)	**Yes** (n = 15)	**No** (N = 33)
Assessment specific to each objective (formulated by the authors)[1]								
Reduction in accidents/ MSDs/pain	**7 (58%)**	**25 (69%)**	**20 (63%)**	**12 (75%)**	**10 (56%)**	**22 (73%)**	**11 (73%)**	**21 (64%)**
Positive assessment	3 (25%)	9 (25%)	5 (16%)	7 (44%)	1 (6%)	11 (37%)	6 (40%)	6 (18%)
No assessment	4 (33%)	16 (44%)	15 (47%)	5 (31%)	9 (50%)	11 (37%)	5 (33%)	15 (45%)
Behaviour changes	**4 (33%)**	**9 (25%)**	**10 (31%)**	**3 (19%)**	**5 (28%)**	**8 (27%)**	**2 (13%)**	**11 (33%)**
Positive assessment	3 (25%)	6 (17%)	6 (19%)	3 (19%)	4 (22%)	5 (17%)	2 (13%)	7 (21%)
No assessment	1 (8%)	3 (8%)	4 (13%)	-	1 (6%)	3 (10%)	-	4 (12%)
Improved physical capacities	**-**	**4 (11%)**	**4 (13%)**	**-**	**-**	**4 (13%)**	**-**	**4 (12%)**
Positive assessment	-	4 (11%)	4 (13%)		-	4 (13%)	-	4 (12%)
Enhanced knowledge	**2 (17%)**	**6 (17%)**	**4 (13%)**	**4 (25%)**	**1 (6%)**	**7 (23%)**	**2 (13%)**	**6 (18%)**
Positive assessment	-	5 (14%)	2 (6%)	3 (19%)	1 (6%)	4 (13%)	1 (7%)	4 (12%)
Mixed assessment	1 (8%)	-	-	1 (6%)	-	1 (3%)	1 (7%)	-
No assessment	1 (8%)	1 (3%)	2 (6%)	-	-	2 (7%)	-	2 (6%)
Reduction in health risk factors	**3 (25%)**	**3 (8%)**	**6 (19%)**	**-**	**4 (22%)**	**2 (7%)**	**4 (27%)**	**2 (6%)**
Positive assessment	1 (8%)	3 (8%)	4 (13%)	-	2 (11%)	2 (7%)	3 (20%)	1 (3%)
Mixed assessment	1 (8%)	-	1 (3%)	-	1 (6%)	-	-	1 (3%)
No assessment	1 (8%)	-	1 (3%)	-	1 (6%)	-	1 (7%)	-
Other[2]	**1 (8%)**	**6 (17%)**	**3 (9%)**	**4 (25%)**	**2 (11%)**	**5 (17%)**	**3 (20%)**	**4 (12%)**
Positive assessment	-	3 (8%)	2 (6%)	1 (6%)	1 (6%)	2 (7%)	-	3 (9%)
Mixed assessment	1 (8%)	-	-	1 (6%)	-	1 (3%)	1 (7%)	-
No assessment	-	3 (8%)	1 (3%)	2 (13%)	1 (6%)	2 (7%)	2 (13%)	1 (3%)

Criterion 1: content adapted to the context; Criterion 2: motor engagement; Criterion 3: practice in a real environment or one representative of work; Criterion 4: changes in work situations

1. *Not mutually exclusive subcategories*

2. *E.g., Evaluate the content of a training program, evaluate the impact of a training program on participation in leisure activities, evaluate the impact of a training program on changes to the work activity, etc.*

Appendix 5 – Presence of Training Quality Criteria for Each Study Selected

Studies selected	Criterion 1	Criterion 2	Criterion 3	Criterion 4
Best (1997)	No	No	No	No
Bewick and Gardner (2000)	Yes	Yes	No	No
Brown et al. (2002)	No	No	No	Yes
Carlton (1987)	No	Yes	No	No
Carrivick et al. (2001)	No	No	No	Yes
Chaffin et al. (1986)	Yes	Yes	Yes	No
Cheng et al. (2009)	Yes	Yes	No	No
Crawford and Weetman-Taylor (1996)	No	No	No	No
Daltroy et al. (1997)	No	Yes	Yes	No
Daynard et al. (2001)	Yes	Yes	Yes	Yes
Dehlin et al. (1981)	No	Yes	No	No
Donchin et al. (1990)	No	Yes	No	No
Fanello et al. (2002)	No	Yes	Yes	No
Feldstein et al. (1993)	No	Yes	Yes	No
Godbey et al. (2002)	No	No	No	No
Guo et al. (1992)	No	Yes	No	No
Hartvigsen et al. (2005)	No	Yes	Yes	No
Hignett and Crumpton (2007)	No	Yes	Yes	No
Hollingdale and Warin (1997)	No	No	No	No
Hultman et al. (1984)	Yes	Yes	Yes	No
Jensen et al. (2006)	No	Yes	Yes	No
Johnsson et al. (2002)	No	Yes	No	No
Jones et al. (1999)	Yes	No	No	Yes
Kraus et al. (2002)	No	No	No	Yes
Kuorinka et al. (1994)	No	Yes	Yes	No
Lavender et al. (2007)	No	Yes	Yes	No
Massy Westropp and Rose (2004)	No	No	No	No
Muto et al. (2008)	No	No	No	Yes
Nygard et al. (1998)	Yes	Yes	Yes	Yes
Ore (2003)	Yes	Yes	No	No
Owen et al. (2002)	No	No	No	Yes
Poosanthanasarn et al. (2005)	No	Yes	Yes	Yes
Reddell et al. (1992)	No	Yes	No	No
Scholey (1983)	Yes	Yes	Yes	No
Scott (1995)	No	No	No	No
Snook et al. (1978)	No	No	No	No
Straker et al. (2004)	Yes	No	No	Yes
St-Vincent et al. (1989)	No	Yes	No	No
Tang (1987)	No	Yes	Yes	No
Van Poppel et al. (1998)	No	Yes	Yes	Yes
Wachs and Parker-Conrad (1989)	No	No	No	No
Warming et al. (2008)	Yes	Yes	No	No
Wickstrom et al. (1993)	No	Yes	No	No
Williams et al. (2002)	No	Yes	No	No
Wood (1987)	Yes	Yes	Yes	Yes
Wright and Haslam (1999)	No	Yes	Yes	Yes
Yassi et al. (2001)	No	Yes	No	Yes
Zadvinskis et al. (2010)	No	No	No	Yes

Appendix 6 – Results of the Evaluation of the Study Design Quality by the Meta-analyses

Studies selected	Cochrane Group				HSE Group		Other	Assessment
	Martimo et al. (2007)		Verbeek et al. (2011)		Haslam et al. (2007)	Clemes et al. (2009)	Hogan et al. (2014)	
	RTC[1]	Cohorts	RTC	Cohorts				
Agruss et al. (2004)	-	-	-	-	Good	Good	-	Other
Asfour et al. (1984a)	-	-	-	-	Poor	Poor	-	Other
Asfour et al. (1984b)	-	-	-	-	Poor	Poor	-	Other
Best (1997)	-	High	Good	-	Average	Average	Average	Other
Bewick and Gardner (2000)	-	-	-	-	NA[2]	-	-	NA
Brown et al. (2002)	-	-	-	-	Poor	Poor	-	Other
Burt et al. (1999)	-	-	-	-	Average	-	-	Other
Carlton (1987)	-	-	-	-	Good	Good	-	Other
Carrivick et al. (2001)	-	-	-	-	High	-	-	High
Chaffin et al. (1986)	-	-	-	-	Poor	Poor	-	Other
Cheng et al. (2009)	-	-	-	Poor	-	-	-	Other
Cornish and Jones (2006)	-	-	-	-	NA	-	-	NA
Crawford and Weetman-Taylor (1996)	-	-	-	-	NA	-	-	NA
Daltroy et al. (1997)	High	-	-	High	High	High	High	High
Daynard et al. (2001)	-	-	-	-	Average	-	-	Other
Dehlin et al. (1981)	-	High	Good	-	-	-	-	Other
Donchin et al. (1990)	-	-	-	-	High	High	-	High
Ellis (1993)	-	-	-	-	NA	-	-	NA
Fanello et al. (2002)	-	High	Good	-	-	-	Average	Other
Feldstein et al. (1993)	-	High	Good	-	Average	Average	Good	Other
Gagon et al. (2003)	-	-	-	-	Poor	Poor	-	Other
Genaidy et al. (1989)	-	-	-	-	Average	Average	-	Other
Genaidy et al. (1990a)	-	-	-	-	Average	Average	-	Other
Genaidy et al. (1990b)	-	-	-	-	Average	Average	-	Other
Genaidy et al. (1991a)	-	-	-	-	Average	Average	-	Other
Genaidy et al. (1991b)	-	-	-	-	Good	Good	-	Other
Genaidy et al. (1994)	-	-	-	-	Average	Average	-	Other
Gladman (1993)	-	-	-	-	NA	-	-	NA
Godbey et al. (2002)	-	-	-	-	Poor	-	-	Other
Gross (1984)	-	-	-	-	Poor	Poor	-	Other
Gundewall et al. (1993)	-	-	-	-	Good	Good	-	Other
Guo et al. (1992)	-	-	-	-	Average	Average	-	Other
Hartvigsen et al. (2005)	-	High	Good	-	High	High	High	High
Hellsing et al. (1993)	-	-	-	-	High	High	-	High
Hignett and Crumpton (2007)	-	-	-	-	NA	-	-	NA
Hollingdale and Warin (1997)	-	-	-	-	NA	-	-	NA
Hultman et al. (1984)	-	-	-	-	Poor	-	-	Other
Jensen et al. (2006)	-	-	-	Poor	-	-	Good	Other
Johnsson et al. (2002)	-	-	-	-	-	-	Average	Other
Jones et al. (1999)	-	-	-	-	Poor	Poor	-	Other

1. *RCT: randomized controlled trial*
2. *Unassigned*

Studies selected	*Cochrane Group*				*HSE Group*		*Other*	Assessment
	Martimo et al. (2007)		**Verbeek et al.** (2011)		**Haslam et al.** (2007)	**Clemes et al.** (2009)	**Hogan et al.** (2014)	
	RTC[1]	Cohorts	RTC	Cohorts				
Kane and Parahoo (1994)	-	-	-	-	NA[2]	-	-	NA
Keijsers et al. (1990)	-	-	-	-	Good	-	-	Other
Kilgariff and Best (1999)	-	-	-	-	NA	-	-	NA
Knapik (1997)	-	-	-	-	Poor	Poor	-	Other
Kraus et al. (2002)	Poor	-	-	High	-	-	-	High
Kuorinka et al. (1994)	-	-	-	-	NA	-	-	NA
Lavender et al. (2007)	-	-	-	Poor	-	-	-	Other
Massy Westropp and Rose (2004)	-	-	-	-	NA	-	-	NA
Muto et al. (2008)	-	-	Good	-	-	-	-	Other
Nygard et al. (1998)	-	-	-	-	Poor	Poor	-	Other
Ore (2003)	-	-	-	-	High	High	-	High
Owen et al. (2002)	-	-	-	-	Average	-	-	Other
Pedersen et al. (2007)	-	-	-	-	-	Average	-	Other
Poosanthanasarn et al. (2005)	-	-	-	-	Good	Good	-	Other
Rabinowitz et al. (1998)	-	-	-	-	Poor	Poor	-	Other
Reddell et al. (1992)	Poor	-	-	Poor	-	-	Good	Other
Resnick et al. (2008)	-	-	-	-	-	Average	-	Other
Saleem et al. (2003)	-	-	-	-	Average	-	-	Other
Scholey (1983)	-	-	-	-	Poor	Poor	-	Other
Scott (1995)	-	-	-	-	NA	-	-	NA
Sharp and Legg (1988)	-	-	-	-	Poor	Poor	-	Other
Snook et al. (1978)	-	-	-	-	NA	-	-	-
Straker et al. (2004)	-	-	-	-	High	-	-	High
Stubbs et al. (1983)	-	-	-	-	Poor	Poor	-	Other
St-Vincent et al. (1989)	-	-	-	-	NA	-	-	NA
Swain et al. (2003)	-	-	-	-	NA	-	-	NA
Tang (1987)	-	-	-	-	NA	-	-	NA
Van Poppel et al. (1998)	High	-	-	High	-	-	High	High
Videman et al. (1989)	-	-	-	-	Average	Average	Poor	Other
Wachs and Parker-Conrad (1989)	-	-	-	-	NA	-	-	NA
Warming et al. (2008)	-	-	-	Poor	-	High	Good	Other
Wickstrom et al. (1993)	-	-	-	-	Poor	-	-	Other
Williams et al. (2002)	-	-	-	-	Poor	Poor	-	Other
Wood (1987)	-	-	-	-	Poor	Poor	-	Other
Wright and Haslam (1999)	-	-	-	-	NA	-	-	NA
Yassi et al. (2001)	Poor	-	-	Poor	Good	-	Good	Other
Zadvinskis et al. (2010)	-	-	Good	-	-	-	-	Other

1. RCT: randomized controlled trial
2. Unassigned