*Wearable Technologies* (2025), 6, e5
doi:10.1017/wtc.2024.28

RESEARCH ARTICLE

# From the idea to the user: a pragmatic multifaceted approach to testing occupational exoskeletons

Christian Di Natali[1], Tommaso Poliero[1], Matteo Sposito[1], Vasco Fanti[1,2], Sergio Leggieri[1] and Darwin G. Caldwell[1]

[1]Department of Advanced Robotics, Istituto Italiano di Tecnologia, XoLab, Genoa, Italy
[2]Department of Informatics, Bioengineering, Robotics and Systems Engineering (DIBRIS), Universita' degli Studi di Genova (UniGe), Genova, Italy
**Corresponding author:** Christian Di Natali; Email: christian.dinatali@iit.it

**Received:** 09 July 2024; **Revised:** 30 October 2024; **Accepted:** 02 December 2024

**Keywords:** industrial exoskeleton; on-site assessment; wearable robotics; fatigue reduction; acceptance; experimental protocol

## Abstract
Assessment of occupational exoskeletons should ideally include longitudinal and multistage studies in real working scenarios to prove their effectiveness and sustainability in real in-field contexts and to help generalize the findings for specific scenarios. This work presents a comprehensive assessment methodology implemented as a multistage experimental campaign for rail industry workers using a back-support exoskeleton (StreamEXO). This work demonstrates that a sector/task-specific exoskeleton developed to address work task-specific requirements generates beneficial performance and user experience results. The experimental work in this paper involves collecting data from nine workers over multiple days of testing. During this testing, workers did not report hindrances to their work operations, with an acceptance rate of 86%. In addition, worker fatigue was reduced by 16.9% as measured through metabolic consumption, and 51% when assessed by perceived effort. This work supports the hypothesis that sector/task-specific exoskeletons when tailored to meet the needs of workers and the work tasks can produce demonstrable benefits in real industrial sectors.

## 1. Introduction
Over 38% of workers in the railway industry (1 million people employed in the EU sector at the end of 2016) (European Commission, 2019) have developed work-related musculoskeletal disorders (WMSDs) during their careers (Schneider et al., 2010), making this job unattractive to young people. Moreover, recent reports have highlighted the aging workforce, especially in Spain, Greece, and Italy, where over 50% of the railway workforce was over 50 years old in 2016. Occupational exoskeletons (OEs) are wearable devices that aim to reduce exposure to risk factors associated with WMSDs and could have a very positive value in this sector.

Exoskeletons aim to reduce the physiological workload on an assisted joint by providing an assistive force/torque. Upper-limb exoskeletons assist the shoulder and/or elbow flexors in performing repetitive tasks. They can also provide an assistive force on the shoulders to help raise the users' hands above their heads (McFarland and Fischer, 2019; Kim et al., 2020). In contrast, back-support devices are made

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial licence (http://creativecommons.org/licenses/by-nc/4.0), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original article is properly cited. The written permission of Cambridge University Press must be obtained prior to any commercial use.

explicitly for manual material handling (MMH) (Kermavnar et al., 2021), where they aim to reduce the compressive load on the spinal discs and the associated risk of injury. They do this by producing/absorbing part of the torques usually generated by back muscles when a worker is lifting a load (De Looze et al., 2016). As the benefits of OEs are still under debate, despite important and valuable short focused laboratory investigations (Kuber et al., 2022), several groups have recently conducted longer assessments of the effectiveness of such exoskeletons in real-world applications (McFarland and Fischer, 2019; Pesenti et al., 2021). Across these studies, different categories of metrics have been found useful to experimentally evaluate the exoskeletons and their impact on the workers (Grazi et al., 2019).

One significant aspect of the assessment of OEs is that existing ergonomic evaluation frameworks (such as the NIOSH lifting equation or the OCRA checklist) (NIOSH, 1981; Occhipinti, 1998) do not yet incorporate OEs within their guidelines. Researchers have therefore often attempted to "fit" potential OE benefits within current standard ergonomic evaluation guidelines (Spada et al., 2018; Di Natali et al., 2021). This has been shown recently in experiments that assess fatigue (Gillette et al., 2022), the cumulative risk of injury (Zelik et al., 2022), and the effect of external loads on the musculoskeletal system (Onofrejová and Balážiková, 2021; van der Have et al., 2023). These studies are almost exclusively laboratory-based, and tasks vary broadly from static to dynamic, that is, a single prolonged repetition to high-frequency movements/repetitions. They are, however, not a true representation of the real work environment, nor do they measure the impact of more prolonged system usage.

OEs usually undergo several stages of testing with human subjects, including validation (prototype testing), evaluation (simulated tasks in a laboratory setting), and field assessment (testing simulations in a real-life environment) (De Bock et al., 2022). Kopp et al. (2022) devised a lab-based framework and protocol to explore tasks that are closer to real work situations. These testing phases are associated with a higher technology readiness level (TRL) (Héder, 2017), particularly from TRL4 to TRL6, which aim to verify the device performance, main functionalities, and user experience before moving into operational environment validation (TRL7) (Kim et al., 2021; Kim et al., 2022). This critical process determines the readiness of a new technology (Parasuraman, 2000) and is required before deploying these technologies in any market. Although industrial interest in exoskeletons is growing, widespread acceptance is hampered by the need for more evidence at these higher TRLs to validate the use of exoskeletons in real settings and with real workers.

Associated with this is the need for a better understanding of the benefits and impacts of these devices. While laboratory settings can help investigate OE's capability to reduce the load on specific body parts in standardized postures and simplified gestures, the device's effectiveness is often use-case-specific. In particular, target use cases might have specific characteristics, for example, specific designs and constraints of the workstations, which introduce further challenges and requirements. Moreover, since several exoskeletons have been developed for particular applications, this does not necessarily guarantee their suitability for different industrial fields. Recent work has also revealed that the assessment of exoskeletons across industrial sectors has mostly focused on the manufacturing industry (Ahmad et al., 2024). Hence, there is an important need to test exoskeletons in other relevant sectors where WMSD is an issue. Thus, to collect vital scientifically verified evidence, it is essential to study each application in dedicated on-site experiments with real workers or reconstructed locations in close collaboration with these end users.

This paper aims to provide a comprehensive assessment, including lab, real-environment, and operational assessments, based on a recent experimental study carried out within the European project STREAM (Smart Tools for Railway works safEty and performAnce iMprovement) (Di Natali et al., 2023). In this study, we perform a multistage investigation of five aspects: planning/brainstorming, worksite inspection, real-environment simulation, lab testing, and assessment in an operational environment. These experimental stages are applied to the construction sector for railway infrastructure maintenance and renewal. This paper describes and discusses the benefits and results of such a systematic evaluation to guide the development process of a sector/task-specific exoskeleton. While reporting the results and experience of assessing an exoskeleton throughout the lab, real environment, and operational setup, key lessons and limitations are reported and discussed.

### *1.1. Background and motivation*

The large-scale adoption of OEs will require a stepwise knowledge-based approach, grounded in the practical requirements and definitions of the needs of the specific industrial sector. This should be followed by careful verification of the system performance through meticulous investigation of work task biomechanics, usability, acceptability, and user experience. This development approach, which falls within the definition of the user-centered design (Monk, 2000), should ideally be conducted in field settings. Observing the workers' daily tasks is paramount to collecting vital information on which and how repetitive and non-repetitive actions are performed (O'Sullivan et al., 2017; Cavatorta et al., 2019). This begins with a clear understanding of the workers' needs in each specific workplace. This requires a complex analysis of use-case-specific effectiveness metrics (Sylla et al., 2014; Otten et al., 2018; Torricelli et al., 2020).

Conducting product-specific validation studies in the field with workers allows stakeholders (workers, health and safety representatives, human resources, and production departments) to evaluate the effectiveness of the devices, providing helpful information and feedback on practical issues (Ármannsdóttir et al., 2020; Torricelli et al., 2020). Indeed, the effectiveness of exoskeletons depends on the specific use cases and industrial sector tests (Crea et al., 2021). Although standardized evaluating processes have not yet been defined (Torricelli and Pons, 2019; Kopp et al., 2022), analyzing the workers' routines and interaction with the environment while quantifying muscle exertion and the occurrence of fatigue constitutes a sound basis for EO evaluation. Verifying the suitability and effectiveness of introducing OEs in work situations is essential.

To the best of our knowledge, there is a lack of comprehensive experimental studies that focus on providing, for each device and target application, an evaluation of the technology performance, the suitability of the device for the target application, and its acceptance to end users (Theurel and Desbrosses, 2019; Kuber et al., 2022). Given the complexity of the device and its functionality, multistage studies, spanning from lab simulation to real working scenarios, are paramount to help generalize the findings collected in specific workplaces to broader application scenarios. Hence, the collection of data from in-field case studies that explore the OEs' effectiveness and long-term useability are vital steps in supporting technology uptake.

## 2. Materials and methods

The experimental protocol presented in this work uses the StreamEXO, a back-support exoskeleton (shown in Figure 1), and its evolution from TRL5 (reproduction of industrial tasks in the laboratory) to TRL7 (assessment performed in real industry and during an operational activity). This will form a complete multistage assessment framework. The exoskeleton has been developed for use by rail workers in a real workplace scenario. The assessment evaluates the exoskeleton's performance, user experience, usability, acceptance, and safety in a real railway working context. This study is aligned with the approach recommended by the ASTM (American Society for Testing and Materials). In particular, ASTM F3474-20 presents a set of options for assessing one or more specific ergonomic parameters concerning human users of exoskeletons.

The five main phases of the experimental study are summarized in Figure 2: demo and brainstorming, inspection, simulation, lab testing, and operational demonstration. During this process, there was a strong emphasis on joined-up operation and co-design with all those involved in the experimental campaign. Due to the different interests and priorities, agreeing on a shared plan was challenging and required several iterations.

The experimental program is detailed in Stage 1: demo, brainstorming, and shared objectives and key performance indicators (KPIs) are identified.

The second stage involves worksite inspection. This is a preparatory activity in which the complexity of the working task is investigated. This requires evaluating all aspects of worker safety, the worksite, and the personnel involved in the experiment. As the device will be tested in a real scenario, all possible issues


***Figure 1.*** *Back-support occupational exoskeleton: StreamEXO.*

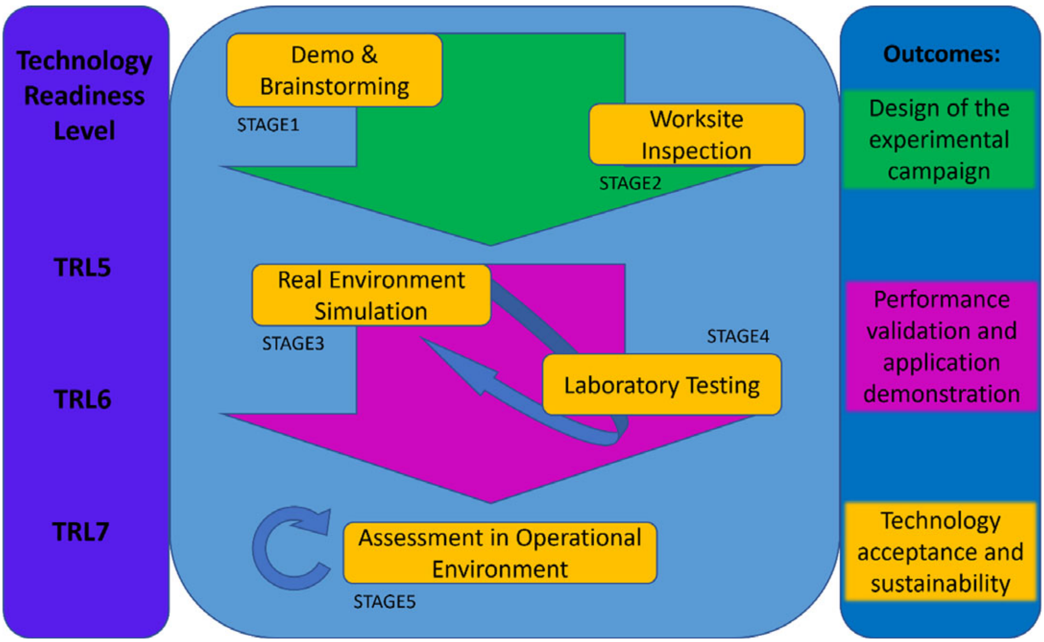


***Figure 2.*** *Methodology of the multistage experimental campaign.*

(e.g., challenges concerning the task execution, team composition, terrain characteristics, environment, rules, directives, work schedules, and the company's task organization) should be identified early to avoid problems during the trials. The analysis of the whole work activity provides data related to timing, duration, production rate, and the company's goals. These are essential to provide an accurate design of the experimental section in Stages 3 and 4 by merging the work characteristics, the KPIs, and the experimental hypothesis.

Stage 3 involves real-environment simulation, and Stage 4 is laboratory testing. Iterative progression between Stages 3 and 4 depends on the KPIs, the nature of the data, and the experimental study's goals. Lab testing supports the validation and demonstration in the real simulated environment, with the collection of relevant data needed to assess the system. Both Stages 3 and 4 are designed to be modular, that is, designing the experimental protocol such that it can be easily repeated with several different workers or with the same worker in several repetitions is crucial. Lab testing aims to support with scientific evidence the performance of the technologies, physiological effects, and potential side effects (motion and muscle pattern modifications). Lab-based results are out of the scope of this work.

The final stage (Stage 5) assesses the operational workplace environment. This stage must address all the technical and safety issues raised during the previous experimental stages. Stage 5 must also allow the workers to carry out their work for a sustained duration (hours, days), in a natural manner with minimal hindrance. Thus, observational and qualitative validation metrics are encouraged.

Due to the nature of the device and the close interaction with the user, several types of metrics and data are required to fully explore and evaluate the exoskeleton's benefits and results fully. Fortunately, the proposed methodology delivered data in many forms: quantitative, qualitative observational, and perceived. The following sections report on the main results concerning the workers' fatigue, task conduction, and the user's experiences. They were evaluated during the work site inspection (Stage 2), real-environment simulation (Stage 3), and operational environment demonstration (Stage 5).

### *2.1. Objectives of the study*

Railway track inspection, maintenance, and renewal work involve many repetitive and strenuous tasks that require considerable physical effort, fatigue, and possible bad postures. Often the operators must lift, replace, move, or carry heavy materials such as sleepers, rails, concrete structures, and wayside objects (telephones, signals, etc.). Other operations can include handling heavy tools, such as moto-cleaves, jacks for track lifting, wrenches, and so forth Further, rail maintenance workers are exposed to severe environmental working conditions (e.g., freezing temperatures in Winter, Summer very high temperatures, wind, rain, dust, working at night, and inside tunnels). Many tasks on the railway can overload the lower back, with an increased risk of WMSDs due to intervertebral overload compression at L5/S1.

A concurrent cause of WMSD during MMH is the high muscular activation of the spinal erectors (NIOSH, 1981), and it is hypothesized that back-support exoskeletons can reduce the risk of injury to a lower and more acceptable rate. Active back-support exoskeletons can support the worker's body by transmitting forces to the shoulders and legs to decrease erector spinae muscle activation and consequently reduce the load on the lower back (Di Natali et al., 2021). However, the benefits cannot easily be generalized, as they are sector-, task-, and person-specific. Hence, there is a need for a dedicated assessment. The technical requirements of the exoskeleton have been co-designed with the direct involvement of end users and stakeholders (Companies, H&S Managers, Project Managers, and Site Managers) (Di Natali et al., 2023).

Concerning requirements, developing any industrially relevant exoskeleton, especially for industries such as rail, needs a demonstration of operation at high TRL, for example, TRL7. This kind of assessment in such an operational environment aims to assess four linked goals:

i. Reduction of overload and fatigue while performing heavy handling activities.
ii. Providing good usability, that is, the device must be safely and effectively used by a pool of workers with a broad range of sizes (small to XX-large) when performing diverse activities.
iii. Providing a positive user experience.
iv. Minimization of physical discomfort and mental stress/fatigue.

In most prior developments and uses of exoskeletons, the exoskeleton system itself has been designed for generic applications, which means that workers can easily do various tasks, much as a worker would. However, exoskeletons are a form of safety equipment, and different requirements and designs of safety equipment are industry-specific; therefore, the working hypothesis is that the design of the sector-specific exoskeleton will be a natural and valuable development for future OEs.

The STREAM project allowed us (i) to investigate the need for custom-made industry-specific exoskeletons or alternatively (ii) to keep the requirements and specifications as broad as possible to suit any industrial sector. The R&D process supported by user-centered design (UCD) (O'Sullivan et al., 2017) has led to the development of a specific exoskeleton for the target sector; in this way, it is expected that the device's applicability to different sectors will be reduced, but it is assumed that an increase in performance and effectiveness will be obtained. The exoskeleton developed within STREAM is an active version called StreamEXO (shown in Figure 1), whose design has been awarded the design award (IfDesign2024). This paper will follow the development and testing of the performance of StreamEXO up to TRL7.

Regarding the development of the exoskeleton, its design has been optimized to maximize safety, comfort, ergonomics, and acceptance: critical workers must not be hindered during routine maintenance operations. By assisting the workers by providing up to 35 Nm of (lower back) torque per motor, the exoskeleton aims to reduce the risk of overload on the lower back. The device features proprioceptive sensors, electric actuation, and advanced bioinspired action-sensitive control modalities to dynamically modulate the amounts of physical assistance provided to the workers (Di Natali et al., 2020). These dedicated control algorithms and assistive strategies have been developed to precisely address sector-, task-, and person-specific needs while ensuring synchronization with the musculoskeletal system and motion patterns (Di Natali et al., 2024b).

Concerning the exoskeleton control modalities, these were analyzed and studied within the first process phases (Stages 1 and 2 of Figure 2), guiding the development of two leading families of assistive strategies. These families targeted the generation of assistive torque in the case of (i) dynamic task execution or (ii) static tasks. For dynamic tasks, the controller regulates movements that require a visible change of the trunk angle, such as during lifting, carrying, and lowering activities. These motions can be smooth and gradual or more rapid depending on the nature of the work involved. For static operations, the dedicated control strategy provides support when the operator is in a fixed (non-moving for several seconds), often bent position while manipulating tools and loads. This static controller compensates for the user's upper body's gravitational load. Technical details are reported in this work (Di Natali et al., 2024b).

StreamEXO aims to help reduce workers' persistent, long-term exposure to WMSD since this is closely correlated with the overload experienced every working day (Fox et al., 2019). Within this work, we have identified correlated indexes (below) that can be used to measure key features and compare these with respect to the primary goal of WMSD reduction. These contributory factors are assessed by dedicated qualitative and quantitative tests that assess the impact and benefits of an exoskeleton. These include the reduction of metabolic costs and the reduction of perceived fatigue (KPI1), the delay in the onset of fatigue (KPI2), the reduction of muscle activity (KPI3), and the reduction of ergonomic risk (KPI4). In addition, another relevant KPI for testing with workers in an operative environment concerns the end user's acceptance of the device (KPI5).

The target KPIs for this multistage study are associated with a dedicated baseline (that reflects the current state) and the target results from using the exoskeleton. KPI4 and KPI5 are reported for completeness but are not addressed in this work. Specific contributions are reported in which achievements concerning these latter KPIs are addressed.

- KPI1: Reduction of fatigue. Measurement of the metabolic cost and perceived exertion while carrying out typical working tasks.
- KPI2: Increase in workers' activity (without fatigue) duration when performing specific heavy tasks. Measurement of the execution time while workers carry out specific heavy tasks.

- KPI3: Reduction of muscular activity. Measurement of the lower-back muscles' activity during regular working activity. Results on the reduction of muscle activity are reported in Di Natali et al. (2024b).
- KPI4: Ergonomic evaluation. Observational measurements and ergonomic evaluation questionnaires (revised NIOSH) evaluated during regular working activity. Results on the reduction of ergonomic risk are reported in Di Natali et al. (2024a).
- KPI5: Acceptance rate of workers during the use of the Exoskeleton in real conditions. Evaluation of user satisfaction while considering several acceptance factors: comfort, assistance, and so forth In addition to the results presented in this work, the supplementary results are reported in Sposito et al. (2024).

### *2.2. Case study*

Several railway construction, inspection, and maintenance tasks are based on the manual handling of heavy materials and objects, such as replacing rails, sleepers, cable ducts, or other parts of the infrastructure. Some of these objects can weigh up to 80–100 kg and must be handled by a team of workers sharing the load. As a result, the individual load ranges from 5 to 30 kg. Additionally, everyday tasks can involve using heavy tools (5–20 kg), often requiring holding the tool for up to 10 min while maintaining a bent-over posture. These activities may be repeated several times per hour, requiring significant physical exertion and subjecting the workers to substantial ergonomic risk (harmful loads on the lumbar spine).

During the planning and subsequent experimental testing shown in Figure 2, several stakeholders were involved. Health and safety managers and workers were consulted from the early trial planning stages onwards, while field testing also involved a range of other managers (production and personnel). Finally, device developers and device performance investigators were also consulted.

The multistage experimental study that forms the focus of this paper involved electric line renewal. This consists of laying concrete ducts along the side of the railway tracks. This activity precedes the laying down of fibers or copper cables for telecommunication, signaling, and supply systems. These ducts are laid in the station yards and along the tracks. They can be part of building new lines, renewals, or expansions.

Rail workers in renewal and maintenance operate in teams of four to six. Team sizes are determined by the need for a sufficient number of people to accomplish the tasks dynamically and quickly move from site to site while potentially covering large distances during a shift. Rail companies form teams according to the tasks, workers' characteristics, and experience. It is recognized that the rail industry lacks gender (and often age) balance in terms of the workforce involved in the field; in the rail industry in Western Europe, less than 10% of the workforce involved in construction/maintenance is female (Eva Heckl, 2010).

In this paper, the work activity explored involves placing concrete ducts flush with the ground, with cables being installed within these ducts. Depending on the purpose, these ducts can weigh 10–50 kg (heavier parts can reach 95 kg but are manipulated with dedicated machines). The placement is performed individually for lighter parts and by a team of two workers for the heavier ones. The workers can use clamps, weighing 7 kg, when handling the heavier cable ducts but often these are not used because they are cumbersome. Generally, the cable laying activity lasts several days. On average, a team of three to four workers can complete 100–200 m daily. When the ducts and cables are installed, they are covered with concrete covers weighing 10–30 kg. Workers are subjected to awkward positions during these activities, such as back and knee bending, which can cause back, shoulder, and neck pain.

A typical working routine involves (1) walking, (2) lifting heavy loads, (3) transporting loads, (4) positioning loads, (5) maintaining incongruous postures, and (6) use of (often heavy) tools such as hammers, blades, and rakes.

The experiment was approved by the Ethical Committee of Liguria (protocol reference number: CER Liguria 001/2019) and complies with the Helsinki Declaration. All the subjects signed a consent form prior to participating, after receiving a full explanation of the experimental procedure.

### 2.3. Data gathering

Based on the selected KPIs, we divided the data acquisition into four categories: Quantitative, Observational, Perceptional, and Qualitative. These data required a specific procedure and methodology to ensure the effectiveness of the data acquisition and avoid any possible artifacts and unwanted signal conditioning of the data. Moreover, since the KPIs in this work are focused on comparing performances or benefits derived from using an exoskeleton, having a verifiable baseline (operation without an exoskeleton) is essential. This section presents the categories of data, and relevant methodological details are reported.

#### 2.3.1. Quantitative assessment

The quantitative data collected included execution time and metabolic consumption.

The execution time was measured during the (i) operational setup tests (Stage 5), (ii) and on-site real-environment simulation (Stage 3). Video recordings of the workers' activities were collected to support postprocessing. For the on-site real-environment testing, the execution time was recorded at the end of the "Fine" positioning of the tenth block of each cycle. During the operational setup testing, the execution time was recorded for each subtask performed, also reporting the number of manipulated blocks.

The fatigue (KPI1) was measured by indirect calorimetry using a method based on the tested subject's respiratory quotient and the rate of $O_2$ utilization. The outcomes were normalized based on the weight of each subject. We measured the metabolic consumption during the on-site real-environment tests. Each worker performed the test wearing a K5 metabolic mask (K5 COSMED Srl, Roma, ITALIA). This equipment provides direct measurement of the exchange of oxygen and carbon dioxide. This returns the metabolic equivalent (MET), which is defined as the amount of oxygen consumed while sitting at rest and is equal to 3.5 ml $O_2$ per kg body weight per minute. The MET is a simple, practical, and easily understood procedure for expressing the energy cost of physical activities. 1 MET is equal to 1.162 W/kg.

#### 2.3.2. Observational assessment

Observational methods are the most common approach in ergonomics to evaluate physical workload to identify hazards, monitor the impacts of ergonomic changes, and assess research outcomes (Takala et al., 2010; Hellig et al., 2019). In addition to objective measurements, working postures were evaluated using the revised NIOSH lifting index (NIOSH, 1981; Waters et al., 2011) and variable lifting index methodologies (Waters et al., 2016), and ISO 11228-1:2021. The NIOSH lifting equation assesses the risk of low-back disorders arising from repeated lifting. Factors relating to the lifting conditions are used to yield a recommended weight limit. Multipliers are based on biomechanical, physiological, psychophysiological, and epidemiological data. The lifting index (LI) and the Variable Lifting Index (VLI) provide the ratio of the actual weight handled to the recommended weight limit. LI values below 1.0 are considered safe for the average population. The greater the index, the greater the risk of low-back injury (Waters 1993, 1994; Marras, 2006). Very importantly, from an ergonomic perspective, there is currently no established procedure to evaluate the benefits provided by an exoskeleton. However, some new methods have been suggested and are being evaluated by researchers (Spada et al., 2018; Di Natali et al., 2021; Zelik et al., 2022).

In this work, observational measurements were applied in all the testing stages, that is, worksite inspection (Stage 2), real-environment simulation (Stage 3), and operational assessment (Stage 4). Furthermore, ergonomic analysis ensured that simulated tests aligned with the real scenario evaluated during the initial worksite inspection. The calculation sheet comprises multiple coefficients required to assess the ergonomic risk of an MMH activity. An example of the parameters selection and relative calculation sheet is reported in Supplementary Section 11.1.

#### *2.3.3. Perceived exertion assessment*

The Borg Rating of Perceived Exertion (RPE) (Borg, 1998a) measures the perceived physical activity intensity which is how hard you feel your body is working. It is based on personal physical sensations during the activity, including increased heart rate, respiration or breathing rate, sweating, and muscle fatigue. Although this is a subjective measure, the exertion rating scale, which is based on five verbal anchors (1-low, 2-normal, 3-moderate, 4-high, and 5-very high) (Dawes et al., 2005), provides a fairly good estimate of the overall fatigue resulting from physical activity (Borg, 1998b). During use, participants are asked to rate their exertion on the 5-level scale considering their feelings of physical stress and fatigue, disregarding any factors such as leg pain or breathlessness, but focusing on the whole feeling of exertion. The number chosen provides a good indication of the intensity of the activity allowing the participant to speed up or slow down movements/activity. Estimating the scale takes only a few seconds and can be self or researcher-administered on a single occasion or multiple times. During the real-environment simulation, the Perceived Exertion was recorded at the end of each working cycle (Tsao et al., 2018). The table used to record the perceived exertion is shown in Supplementary Section 11.2.

In addition, other evaluations were conducted by applying the Local Perceived Discomfort (LPD) scale (Van der Grinten, 1991) using a Borg scale assessment for different body sections to evaluate the discomfort of using the exoskeleton (Borg, 1998b). A body map consisting of eight regions of the anterior and posterior body was presented to the subjects (Van der Grinten, 1991; Bosch et al., 2016). Subjects were asked to rate their discomfort in the regions identified on a 10-point scale (ranging from 0 no discomfort to 10 extreme discomforts, almost maximum). This Rating of Perceived Exertion and the LPD has been used for the real-environment simulation and the operational experimental test runs.

#### *2.3.4. Qualitative assessment*

At the end of each experimental test day, each subject completed an interview to assess their experience with the exoskeleton. This questionnaire covers the following topics: assistance provided by the exoskeleton, comfort, stability of the device on the user, usability, and device acceptance. This questionnaire has been validated (Sposito et al., 2021) and used with minor changes throughout all the experimental stages. Since the workers used the exoskeleton for several days (3 days for each worker) during the real-environment simulations and the operational worksite testing (2–4 days for each worker), they had an opportunity to accumulate considerable experience using the device. Hence, it was also possible to study trends and the evolution of intrasubject data could be evaluated as the experience of using the exoskeleton increased. In this work, we reported the final impression of each worker after at least four and a half hours of use. The interviews were performed with each worker individually to avoid possible interference and bias introduced by colleagues.

Each category comprises multiple questions with a fully anchored 7-point Likert-type scale reported in Supplementary Section 11.3.

## 3. Results

### *3.1. Stage 2: work site inspection*

The worksite inspection is fundamental to evaluating the feasibility of the experimental campaign. This stage incorporates task analysis, which evaluates all the subtasks and activities. This operation allows the evaluators to understand the process associated with the activities and identify minor sub-activities, particular movements, loads, tools involved, and so forth Then, a sequence of subtasks is carefully determined and selected based on which one scored highest on the ergonomic risk assessment.

To calculate the risk index for operations occurring in a typical working day, experts in ergonomic evaluations used the revised NIOSH lifting equation (Waters et al., 2011) as this is aimed at MMH activities.

A simplified "typical working day" involves the following activities. Figure 3 shows the six phases of electric line renewal where new cable ducts are placed alongside the track. The tasks consist of placing approximately 200 conduits of 50 kg each for a covered distance of 100 m daily. In addition, once the duct is laid, a cable bundle is settled into the conduit, requiring about 100 lifts of cables every meter (on-site measurements show an average bundle weight of 24 kg). Finally, covers are placed on top of the ducts (cover weight 30 kg). This electric line renewal task is performed by two workers and involves transporting and laying down the concrete cable ducts ("Gross positioning"), settling them in the trench ("Fine positioning"), positioning the cables, and covering the conduit.

The NIOSH VLI (Waters et al., 2016) combined with ISO 11228-1:2021 applied to the target task results in an LI of 3.36 and 4.2 if workers are younger, or older than 45 years old, respectively. Based on this preliminary analysis, the tasks with the greatest ergonomic risk are Phases 2 and 3 in Figure 3, defined as "Gross positioning" and "Fine positioning" of the cable conduits. Stage 3 is designed to assess the benefit of the exoskeleton while performing "Gross" and "Fine" positioning.

## *3.2. Stage 3: real-environment simulation*

### *3.2.1. Experimental protocol*

The real-environment test run (simulation at TRL6 – real work but in a controlled environment) aims to evaluate the performance of StreamEXO (Figure 1) against KPI1 and KPI2, for example, reduction of fatigue and increase in workers' activity, respectively (Section 2.1). As noted above, in electric line renewal, Task Phases 2 and 3 (Figure 3) are the subtasks with the most significant impact on worker health.

The test run was performed at an Italian rail worksite near Asti train station, under the guidance of an Italian infrastructure manager (RFI). The system integrator company (MERMEC STE) hosted our team and supervised the testing with their workers at the work site. A team of five workers was recruited for testing in Stage 3 (age 49.8 ± 4.5 years; height 175.7 ± 4.7 cm; weight 86.6 ± 17.6 kg.). The test was performed in accordance with the experimental protocol approved by the Ethics Committee of Liguria, Italy (protocol number: 001/2019) and complied with the Helsinki

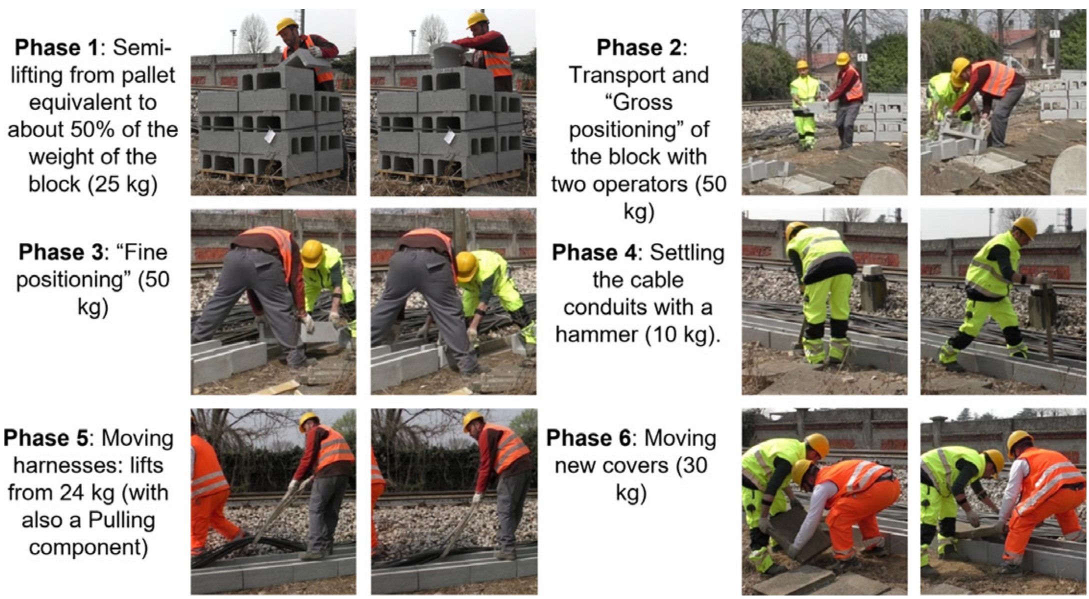


***Figure 3.*** *Cable conduits (duct module V3134 weighing 50 kg) placement activity during a "typical working day."*

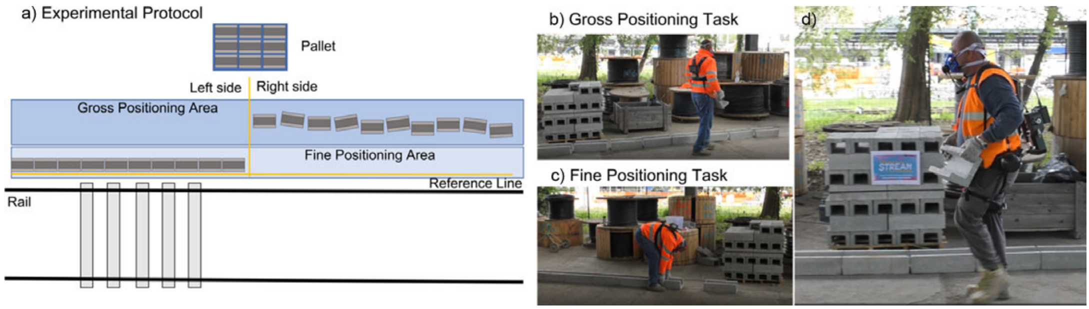


***Figure 4.*** *(a) Experimental protocol for the simulation of cable duct (duct module V318 of 20 kg) placement in a controlled on-site environment. Main test phases: (b) gross and (c) fine positioning. (d) Tests performed by the workers wearing the StreamEXO and the metabolic consumption measurement system (COSMED K5).*

Declaration. The experimental protocol, shown in Figure 4a, was designed to assess the two critical activities ("Gross" and "Fine" positioning) cyclically. The test duration was set to 90 min for each of the two test modalities (with the exoskeleton [EXO] and without the exoskeleton [baseline: NOE]). The test was conducted using 20 kg blocks to allow workers to carry out blocks independently. The NIOSH LI of the "Gross" positioning for the test in Stage 3 (Supplementary Figure 10 in the supplementary sheet [NIOSH calculation sheet]) is comparable with the one calculated in Stage 2 (Section 3.1).

Before the test, each subject underwent theoretical and practical training to fully understand the experimental procedure, including familiarization with the use of the exoskeleton and systems. The familiarization lasted ~30 min for each worker. The worker was required to pick and carry a duct from the top of a pallet and then deploy it near the starting point of the duct placement line (shown in Figure 4b as a reference line). The "Gross" positioning places the ducts ~50–60 cm from the reference line. The workers repeat this subtask 10 times to create a line of ducts ~5 m long. Then, the worker performed the "Fine" positioning subtask. Each worker positions his legs spread apart (a wide stance) and his back bent to place all 10 cable ducts precisely along the reference line (shown in Figure 4c). Each worker lifted the first cable duct closer to the vertical line (between the right and left parts of the testing scenario) and moved it 1 m to cross the vertical line. Then, the worker lifts the second duct by 2 m, lowering it close to the first cable duct. This activity was repeated 10 times until all the ducts were moved from the right to the left side of the test area. Finally, the last duct was carried 10 m. Then, the "Fine" positioning is repeated on the left side of the testing area. Once the worker has completed both phases, they start the test again with the "Gross" positioning of the same 10 cable ducts. Each test cycle (Repetition), comprising "Gross" and "Fine" positioning, is repeated for 90 min for each testing modality.

At the end of the first 90 min test section, the worker rested for 30 min before starting the second modality of the test for the same duration. The testing was repeated three times on three separate days. A randomized order of the test modality (with and without exoskeleton) was chosen at the beginning of the experimental tests. Quantitative, perceived (using the questionnaire in Supplementary Section 11.2), and observational (using the questionnaire in Section 11.1) results reported in Di Natali et al. (2024a) and qualitative data (using the questionnaire in Supplementary Section 11.3) presented in Sposito et al. (2024) were gathered to assess the target KPIs.

### *3.2.2. Real-environment simulation: results*

The reduction in fatigue (KPI1) is evaluated using a quantitative method, based on an indirect measurement of the metabolic cost (Figure 4d), which is derived from the direct measurement of oxygen and

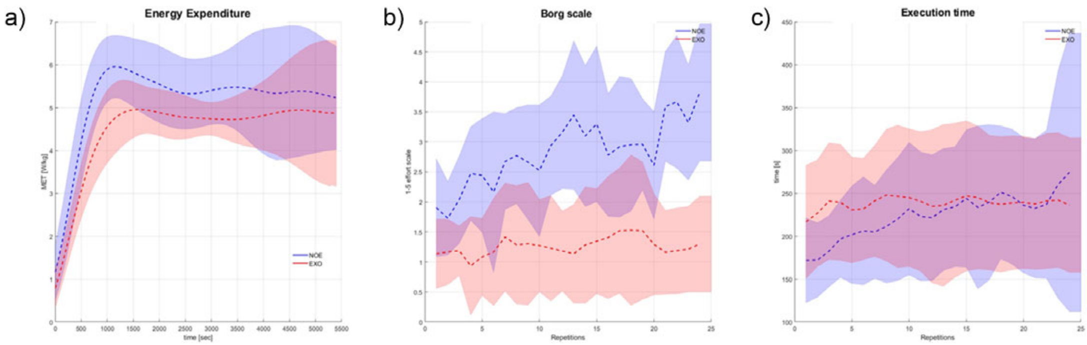


***Figure 5.*** *These figures show mean (dashed line) and standard deviation spreads. (a) Comparison of Baseline (NOE) versus EXO modality of the metabolic consumption trends of all five workers averaged over the three test days (90 min with the exoskeleton and 90 min without the exoskeleton). Numerical results are normalized for each subject's mass. (b) Borg scale on the overall perceived exertion level. (c) Execution time for each repetition of the working cycle under evaluation. Shaded regions represent the standard deviation of each trend.*

carbon dioxide exchanged volume and ventilation (DeBlois et al., 2021). During testing, the metabolic consumption was measured throughout the full 90 min trials, both without the exoskeleton (baseline – NOE) and with the exoskeleton (EXO).

Figure 5a shows the energy cost for typical working activities using the two configurations (NOE and EXO). The average metabolic expenditure, normalized for the subjects' weight, associated with the working activity, and measured over three tests for each of the five workers in the NOE configuration is 5.19 ± 0.9 W/kg (4.47 MET) with a peak of 5.97 ± 0.96 W/kg (5.14 MET). For the EXO configuration, under the same conditions, the metabolic expenditure is 4.47 ± 0.9 W/kg (3.85 MET) on average with a peak of 4.96 ± 0.91 W/kg (4.26 MET). Hence, the average reduction in metabolic cost is 13.8%, while the peak reduction is 16.9%.

A further indicator correlated to fatigue is KPI2 (increase in workers' activity), which aims to evaluate any delay in the onset of fatigue while carrying out heavy job activities. This indicator combines two measurements: the perceived exertion and the measurement of the execution time. The perceived exertion, measured using the five verbal anchors associated with the Borg scale (Dawes et al., 2005), is measured by asking the workers to quantify their effort while carrying out the working task (with and without the exoskeleton) to evaluate possible differences in the effort required to accomplish the tasks. The execution time was measured at the end of each work cycle. An important consideration associated with the execution time is whether the exoskeleton hinders the workers, as the natural execution of the task would be affected, and more time would be required to accomplish it. If the assistance provided by the exoskeleton effectively reduces fatigue, the workers (theoretically) can continue working without the need to take a break or reduce their speed. This could increase productivity, but any potential increase in work period or productivity is not within the scope of this study.

Figure 5b shows the perceived exertion level on the Borg scale for each working cycle, averaging over the five workers and their three task repetitions. Within the 90 min tests, on average, workers performed ~25 working cycles. The average perceived exertion level without the exoskeleton is 2.6 ± 0.5 on the Borg scale. This means that over the 90 min, the task is perceived as almost moderate, while toward the end of the test, the value rises to a maximum of 3.7, meaning nearly high. The average perceived exertion level when using the exoskeleton (EXO) is 1.25 ± 0.14 on the Borg scale, with a final value, after 25 repetitions of 1.32. The task is perceived slightly above low for the entire 90 min test, and there is almost no perceived increase in fatigue at the end compared to the start. The percentage reduction in perceived exertion level based on the average scores after 90 min of testing is 51%.

Regarding the execution time (shown in Figure 5c), the average over the 90 min in the baseline (NOE) modality is 226.8 ± 61.2 s against the 239.7 ± 47.4 of the EXO. The two average execution times differ by ~13 s between the two modalities, corresponding to a 5.7% reduction in production when using the exoskeleton. However, the results show that the EXO reduces the time fluctuations, allowing the worker to work at a more constant speed with a reduced variation (work cycle time execution is more consistent). Moreover, Figure 5c shows that although the workers in the baseline modality start working at a very high speed (~180 s per cycle) after 10 repetitions (~30 min from the beginning of the test), the NOE execution time has dropped below its mean execution time (10th cycle execution time with NOE is 232 s). Around the 20th cycle, it is visible how the NOE execution time strongly increases by more than 50 s in the last five cycles, reaching 285 s as the average execution time. This increase in execution time, and particularly the rapid increase in the last five cycles, essentially sets the 90 min work period before there is a break. With the exoskeleton, as already noted, this increase does not occur. Further, the rapid increase in the latter cycles when not using the exoskeleton, that is, the period of greatest fatigue, is also the time most associated with the risk of injury. As this increase is not observed when using the exoskeleton this could be seen as a very positive impact, although it cannot yet be confirmed that it would definitely reduce injury risk.

### 3.3. Stage 5: assessment in an operational environment

#### 3.3.1. Experimental protocol

The assessment in an in-field operational environment aims to evaluate the use of the exoskeleton under real working conditions (equivalent to TRL7). Bulky equipment for measurements (e.g., metabolic consumption systems) is not suitable at this experimental stage to avoid bias and adversely affect the user experience. This testing stage aimed to quantify specific KPIs and demonstrate the feasibility of using the exoskeleton in a harsh (heavy-duty) rail worksite. In Stage 5, we measure the reduction of perceived effort (KPI1), the execution and task conduction time (KPI2), and the acceptance rate (KPI5).

A team of workers conducted the electric line renewal, following their natural workflow rhythm and the flow of various activities, deciding how and when to move between the multiple subtasks. This is less tightly controlled than in the lab environment (Stage 3) but did allow for more realistic data collection.

The experiment was performed on-site in an Italian rail worksite (Gallarate train station close to Milan) under the guidance of an Italian infrastructure management company (RFI). The workers were employed by a system integrator company (MERMEC STE). Working in teams of four workers (age 46.7 ± 9.5 years; height 180.5 ± 3.3 cm; weight 88.0 ± 3.5 kg.) underwent theoretical and practical training and familiarization with the exoskeleton devices and user interfaces (30 min for each worker). This was undertaken a few days before the test runs.

The operation and maintenance of railway infrastructures are complex, strictly regulated by government legislation, and involve working in close cooperation with all the involved stakeholders. All were fully consulted on safety, the worksite tasks flow, and accomplishing these activities. Before starting testing, the infrastructure manager took "possession" of the railway line to avoid possible train traffic during the morning hours (08.00–14.00).

The team of workers (Figure 6) transported and deployed the cable ducting blocks, which had weights of 30 kg, to achieve the "Gross" and "Fine" positioning. Two workers per time (Wk1 and Wk2) wore two StreamEXOs, allowing them to work in parallel. Wk1 was responsible for transporting the conduits from the maintenance train to the site (an average distance of 5–10 m). Wk2 positioned the conduits in the trench (previously prepared by two additional colleagues who were not wearing an exoskeleton), and then Wk2 settled the conduits within this channel.

The workers were asked to work as naturally as possible, as they would do without the exoskeleton. The testing was repeated over multiple sessions to ensure good repeatability, comfort, and familiarity: four half days of testing (4 h per day) for each of the four workers (two workers per "shift" were wearing the exoskeletons) for a total of 32 hr of testing.


***Figure 6.*** *Testing of the in-field operational setup during which conduit module TT3135 weighing 30 kg was handled to perform (a-b) Gross Positioning (GP) task, and (c-d) Fine Positioning (FP) task.*

#### *3.3.2. Operational environment: task conduction*

Full-track safety precautions were observed. In the in-field operational tests at TRL7, we observed eight subtasks:

1. The rail **T**ruck **U**nloading: from truck to ground level and gross position (TU)
2. **F**ine **P**ositioning of the cable duct into the trench (FP)
3. **U**npalletizing cable ducts from the train **T**ruck (UT)
4. The **G**ross **P**ositioning of the cable duct along the trench (GP)
5. **T**rench **C**leaning and leveling using a rake (TC)
6. Fine positioning and **S**ettling of the cable duct with a **H**ammer (SH)
7. Cable duct **C**over **H**andling and placement (CH)
8. Performance of **O**ther **T**asks, such as driving, using the disk saw (OT).

Figure 7a shows the subtasks on one test day. The two columns on the left side of the chart report the sequence of subtasks conducted during the 4 hr of test for the first and second workers starting from the bottom. Orange shows the combined worker duration of each subtask in minutes, while blue indicates the number of blocks manipulated by the two workers (cable ducts and covers).

Figure 7b,c reports the total number of concrete blocks manipulated during the test and each subtask's net duration, respectively. In particular, Figure 7b shows the total number of blocks and their percentage of the total. The two workers handled more than 1,100 blocks (almost 25 tonnes) during the four half-day tests, where 48% were covers and the rest were cable ducts. The covers weigh 15 kg, and the conduits are 30 kg (conduits module TT3135). The two subtasks that do not require holding weights were the performance of Trench Cleaning (TC) and Other Tasks (OT).

Figure 7c reports the duration in minutes for each subtask. The transportation of the ducts from the ground or the truck (TU, UT, and GP) represents 26% of the total time and is characterized by the dynamic

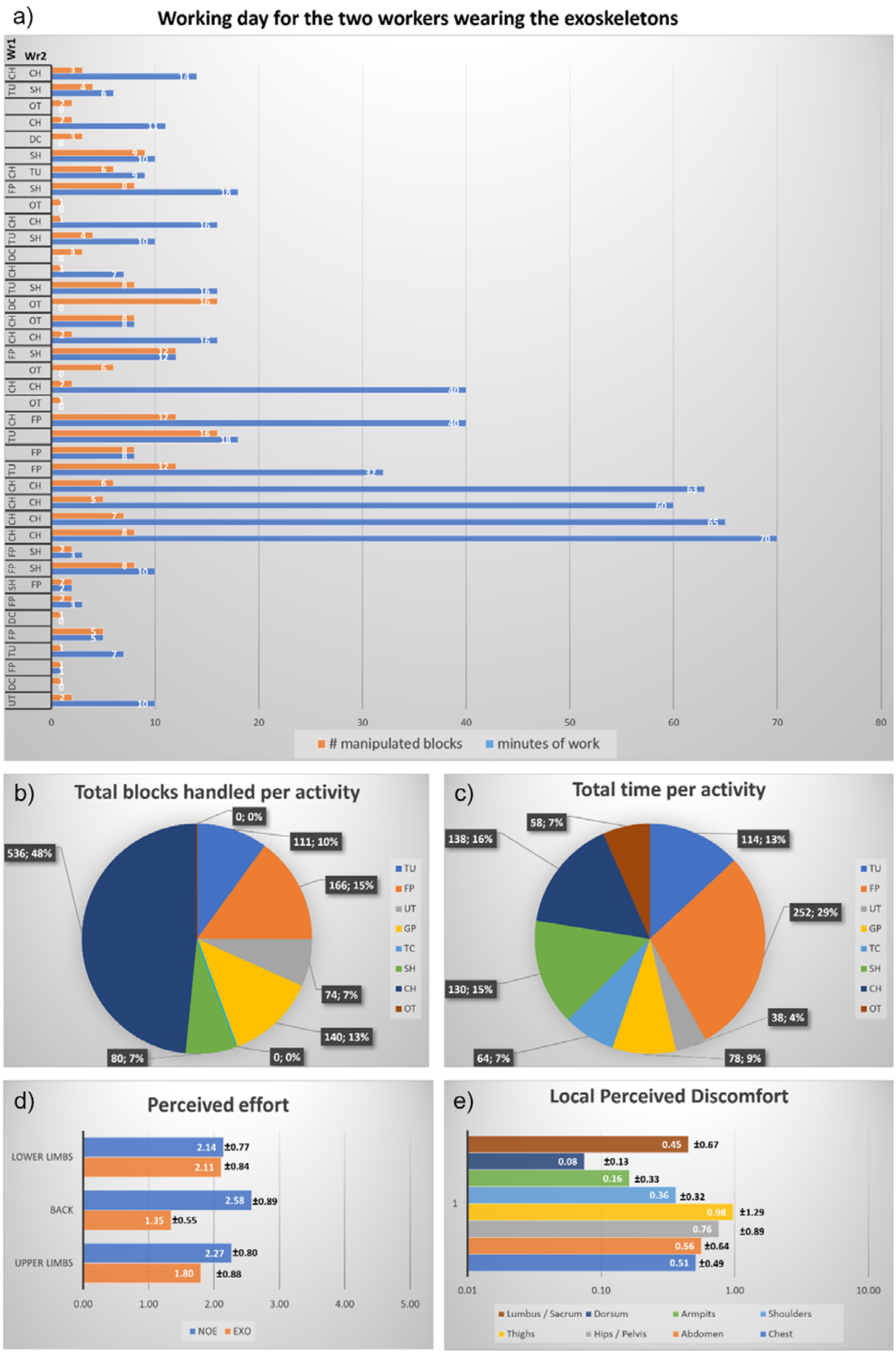


***Figure 7.*** *(a) A typical working day for both workers while performing TU, FP, UT, GP, SH, CH, and OT. The blue line indicates the number of blocks moved, and the orange indicates the time dedicated to each task. (b) The total number of covers and cable ducts manipulated during the experimental activity. (c) The total time for each of the subtasks. (d) Perceived effort at the upper limbs, back, and lower limbs averaged for the main activities of "gross positioning" (lifting, transportation, and lowering) and "fine positioning." (e) Local perceived discomfort measured on eight body regions represented on a logarithmic scale.*

(and highly physically demanding) activities of lifting, transportation, and lowering (Figure 6a). These tasks do not require precision; therefore, these subtasks handle a large number of blocks in a relatively short time, making this task even more risky from an ergonomic perspective. The settling of the cable ducts (SH) into the trench and its subsequent fine positioning (FP) make up 44% of the combined total time. These activities are associated with a static and often awkward posture, where the lifting activity is performed to place the conduit into the trench (Figure 6d). A hammer is then used to level the ducts and ensure good positioning. These activities require precision and several duct adjustments to locate the duct in the trench accurately (Figure 6c); thus, a longer duration is justified. The placement of the covers does not require precise movements, and the relatively lower weight of the covers allows workers to deploy about half of the material handled in a sixth of the total time. Finally, collateral tasks such as Trench Cleaning (TC) and Other Tasks (OT) represent 7% of the entire task duration. Very importantly, these collateral subtasks are fragmented during the whole execution, and this becomes very relevant for the acceptance of the exoskeleton, as the workers must perform these subtasks naturally while wearing the exoskeleton, without removing it.

#### *3.3.3. Operational environment: perceived exertion*

After their experience with the exoskeleton, the nine workers from the two trials completed a questionnaire on their perceived effort. These workers represent the pool of workers from the tests in the simulated and operational environments (five for the relevant environment simulation and four for the operational site). At the end of their test, the subjects were asked to rate the effort required to manipulate the cable conduits and covers. This was separated into three prominent families of subtasks: transportation from and on the truck (TU and UT), "Gross" and "Fine" positioning (that also includes SH) of conduits and covers. Figure 7d reports the average effort perceived (Borg Scale) on these tasks for the lower limbs, back, and upper limbs. From the results, it can be seen that the perceived reduction in effort is 1.6% for the lower limbs, 47.7% for the back, and 20.7% for the upper limbs.

#### *3.3.4. Operational environment: discomfort*

Discomfort in the back (Lumbus/Sacrum and Dorsum), legs (Thighs and Hips/Pelvis), Chest, Abdomen, and Upper body (Shoulders and Armpit) were measured using the LPD scale. At the end of the task, subjects were asked to rate their discomfort in the regions identified on a 10-point scale (ranging from 0 [no discomfort] to 10 [extreme discomfort, almost maximum]). Figure 7e shows the results of the discomfort analysis collected on the workers, each of whom had at least 4.5 hr experience of using the exoskeleton. This included workers who tested the device in the simulated and operational environments (Stages 3 and 5 in Figure 2). The results show a maximum discomfort of about 1 on the Thighs and Hips/Pelvis, corresponding to a "very weak." A barely noticeable discomfort (0.5 scores, "extremely weak") was evaluated on the Lumbar/Sacrum, Abdomen, and Chest, while the rest are below this value. Therefore, the result of the perceived discomfort, after approximately 80 hr of testing from nine workers, suggests that StreamEXO is comfortable, but obviously, further testing will be required in the future.

#### *3.3.5. Operational environment: acceptance*

KPI5 aims to evaluate the acceptance rate of the exoskeleton by measuring the user experience. These qualitative data are acquired by interviewing the nine workers (Figure 2). A custom-validated questionnaire (Sposito et al., 2021) of five categories evaluated the workers' experiences. Each class comprises multiple questions as reported in Supplementary Table 2, with a fully anchored 7-point Likert-type scale: Assistance, Comfort, Stability, Usability, and Acceptance. The questionnaire was administered to each worker, asking them to rate their perception of the continuous use of the exoskeleton in their specific working context, based on their recent experience.

Figure 8 shows the five categories' aggregated results, with their average median score: Assistance 6.75, Comfort 6, Stability 5, Usability 7, and Acceptance 6. The scores associated with using the

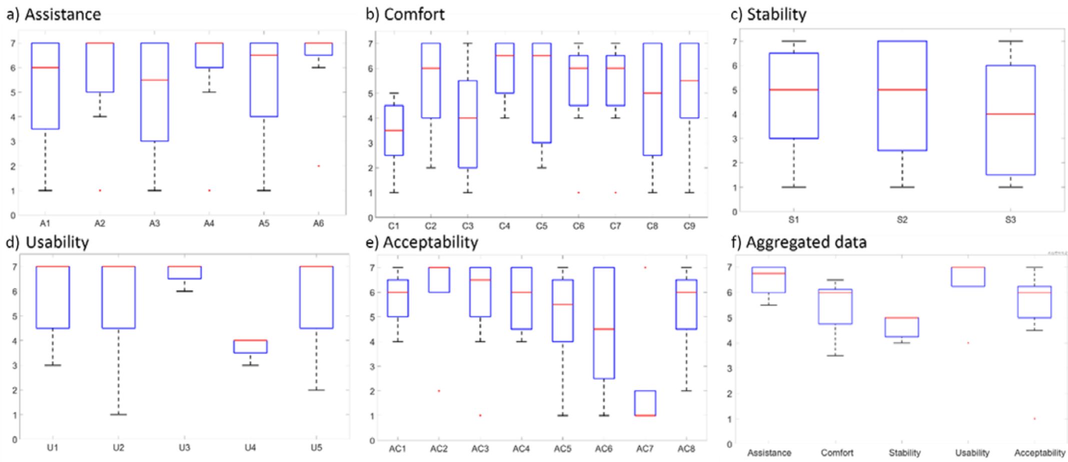


***Figure 8.*** *Figures show the median (red line); the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points, incorporating max and min. Distributions of (a) assistance, (b) comfort, (c) stability, (d) usability, (e) acceptability, and (f) the aggregated data analyzed as a median value over the five categories. Red lines represent the mean score of each question or group of questions.*

exoskeleton resulted in an overall median score of 6.15 on a 7-point Likert-like scale, with a total acceptance rate for StreamEXO of 88%.

The results shown in Figure 8 are obtained with a descriptive statistical analysis. The data show that the overall average scores in the *Assistance* category are all between 5.5 and 7, with a median of 6.75, indicating that workers appreciated the reduction of strain on their back, the amount of assistive force generated, and the unconstrained back and leg movements. The category *Comfort* has a median of 6, showing that the design of the exoskeleton is comfortable, that the transmission of the assistive force is suitably distributed, and that there is no discomfort generated in the legs (C6), chest (C4) and belt area (C5). However, workers reported a median score of 4 for C3 (harness too tight). This could be improved by loosening the straps, but this could negatively impact stability, and this is one area for further development in the future. Similarly, a rating of 3.5 for C1 (the exoskeleton weight is fine) suggests that the weight of the exoskeleton has a negative impact, and future efforts will focus on further reducing this load. At the same time, it was noted that although the mass was a problem, the overall load was well distributed with a score of 6 (C2). Finally, the thermal dissipation is adequate, with a mean score of 5.5 for C7–C9. *Stability* is a sub-category of *Comfort*, and any score below 3.5 could be a significant cause of discomfort. Thanks to the exoskeleton design and considering the workers' experience gained during the days of testing, once the exoskeleton is tightly secured, it is not easy to loosen, and the stability of the device is maintained throughout usage (median of 5). Moreover, workers underlined the need to improve the balance of tightness to provide stability without impacting the perceived comfort. The category *Usability* has a median of 7 for the five questions. The lowest value is for question U4, which concerns the procedure to turn on and off the exoskeleton. During the experiments, the investigators turned off the exoskeleton because of the need to properly save and transmit recorded data from the exoskeleton to external peripherals. This meant the users had no independent control and this was not positively received, however, in a final system, this functional control could easily be given to the users. Finally, *Acceptance* evaluates the workers' overall perception of using the exoskeleton. This category has a median of 6 when considering all questions; in particular, the exoskeleton met their expectations (AC1) and is suited to the workers' tasks (AC2). Workers also reported that they did not feel hindered during the task (AC3) and that they would regularly use it if available on the market (AC4). Question AC7 (I would use the exoskeleton for all my tasks) received a score of 1 since not all the team's tasks concerned MMH. Therefore, the workers would prefer to avoid wearing the exoskeleton if the task does not require assistance.

## 4. Discussion

The organization of the pilot tests required management and solution of several operational rail-related issues. Safety is first and most important and was fully guided by rail industry safety standards for working on real sites with actively moving traffic. This required close collaboration with the safety managers of the workers and with expert ergonomists to achieve the most realistic results during the testing.

Multistage studies are essential to prove the efficacy of OEs and to support their future uptake in the real industrial world. The active tailored back-support OE, StreamEXO, has been developed to meet railway industry needs and worksite challenges. Positive results emerged from these experimental studies. The data shows that StreamEXO effectively reduces workers' fatigue during testing for several days in a real working context. At the same time, there was also broad acceptance by all the workers.

In particular, when converting the results obtained for the metabolic consumption in MET, we noticed that the activity without the exoskeleton is classifiable as an activity of almost vigorous intensity (defined as higher than ~6 W/kg), such as cycling. In contrast, using the exoskeleton significantly reduces metabolic consumption to below 5 W/kg (moderate activity), which is comparable to playing table tennis.

In terms of execution time, the results show an increasing time period to complete the task with the NOE condition, while in the EXO condition, although the initial cycle times are longer than for NOE the execution time remains almost constant over the full 90 min of the testing. The NOE's execution time increases due to increased fatigue, forcing the workers to slow down their work rate. The increment of execution time has the presence of spikes on the 10th, 15th, 18th, and 24th cycles. This is associated with breaks occurring when workers have taken to recover from their fatigue. In contrast, peaks or increases in the EXO execution time are not so evident. Thus, there is an underlying increase in fatigue related to the execution of the task for workers who do not wear the exoskeleton, which is supported by Figure 5b. In addition, this "fatigue time" may correlate with periods that are particularly prone to injury. We could suppose that workers wearing exoskeletons do not feel the same rise in tiredness in the first 90 min of the test, as there is no slowdown in their activity. Therefore, we can consider that the first break requested by the worker due to fatigue would occur at the end of the test (at the 90th min of the test), while for the NOE condition, the worker requires a break on average around the 18th cycle (approximately at the 72th min). Thus, the exoskeleton delays the onset of fatigue by at least 25%. Nevertheless, although this is a promising hypothesis, this study does not yet have sufficient data to confirm this, and future work will develop more extended tests in which we could verify the production rate and the onset of fatigue. Moreover, the test conducted in an operational environment allowed us to verify if the exoskeleton would hinder regular activity. Our intent was not to show an increment in productivity.

Concerning discomfort, it is interesting to compare these results with those from Sposito et al. (2024), which reports on data gathered from five workers performing the test in the simulated environment only. The results shown in Figure 7e combine data from the simulated environment with that gathered during the operational assessment. Sposito reported a discomfort felt at the thighs with a median score of 5 (results measured at Stage 3 in Sposito et al., 2024), which reduced to ~1 (Figure 7e) when combined with the data gathered in Stage 5. This suggests that extended exposure in the real environment reduces the perceived discomfort due to a progressive adaptation to the exoskeleton.

A similar consideration can be made by comparing the acceptance rate measured in Sposito et al. (2024), where 3 days of tests and their trends are reported, with the results obtained in this work, particularly with the aggregated data shown in Figure 8f. In addition, these results are, in one case, the average of the aggregated results of the 3 days of the test in Stage 3 (Sposito et al.), and in the second case, the aggregated results over the 4 days of the test in Stage 5. The values obtained for StreamEXO are sufficiently high to suggest that there is good worker acceptance (from 4.8 to 6). Assistance, Comfort, and Usability are confirmed to be good for the task and sector. The scores in these three constructs in this paper are slightly more positive when compared with the results of Sposito et al. (Assistance rises from 6.4 to 6.75 in this work, Comfort from 4.8 to 6, and Usability from 6.6 to 7). Finally, the results with the lowest user acceptance relate to Stability with 5. These values are similar to Sposito et al. and confirm an underlying need to focus on improving the Stability of the user. While the increases in Assistance and Usability are marginal and already show high acceptance, the increase in comfort is significant (over one full point) and changes from a relatively middle level to a high value.

Further, when comparing sector/task-specific and generic exoskeletons, Fanti et al. (2024) presented a comparison between StreamEXO and two other exoskeletons (Laevo, a passive commercial exoskeleton, and XoTrunk, another active research platform exoskeleton developed at IIT) conducted during Stage 4. The results suggest that the development of sector- and task-specific exoskeletons strongly improves the benefits provided to the user when by reducing muscle activity in the ES and Biceps and Rectus Femoris. This generally suggests that there is growing evidence supporting the beneficial use of exoskeleton in human–exoskeleton physical interactions. This is an important trend, but additional results will continue to be needed for some time before there is definitive proof.

By using a UCD approach, it has been shown that the StreamEXO was able to develop and adapt, for the Rail sector, previous versions of back-support exoskeletons fabricated at the Istituto Italiano di Tecnologia (Sposito et al., 2020). This approach together with close collaborations with project stakeholders, delivered an exoskeleton that has been assessed, accepted, and positively used by workers in an in-field operational worksite environment. Further, more extensive testing will be required to fully confirm these findings, however, this work already suggests some very positive features.

Key lessons include:

- Nontask-specific or generic exoskeletons (or exoskeletons developed having in mind a different industrial sector/task) may not be suitable for use for all tasks, and to achieve good performance in terms of efficacy and acceptance, exoskeletons may need to be sector/task-specific, being tailored in the design and control strategies to meet the work task requirements and needs of the industrial sector since there can be significant work task differences. (e.g., some variable or dynamic tasks could be incompatible with certain passive exoskeletons) (Di Natali et al., 2024b; Fanti et al., 2024).
- The designer should consider several sector-specific aspects (Di Natali et al., 2023): device encumbrance, weight, body size, the target environment, integration with the PPE (personal protective equipment), and so forth.
- Workers' size variability is an important consideration. An exoskeleton must fit appropriately on the user to prevent discomfort and loss of efficacy (Sposito et al., 2021). The particular design developed for StreamEXO presents a continuous size adjustment system that enables the best fit of users for active industrial exoskeletons (Di Natali et al., 2022).

The UCD approach enables us to refine the design of the exoskeletons moving through the assessment stages detailed in Figure 2. Figure 9 shows the three developed versions of the STREAM project. The StreamEXO design has been improved thanks to the feedback from the users during the assessment phases presented in this work. The main improvements throughout the process strengthened the structure, reduced mechanical failures, improved the comfort, and eased the wearing. In addition, the aesthetic aspect was also taken into consideration.

Concerning the usability of the exoskeleton, we made a simple and slim user interface that allows the worker to select autonomously between different control modes. Workers are able to select the dynamic assistance modality during the "Gross" positioning tasks that require lifting, carrying, and lowering activities. In contrast, in the "Fine" positioning operation, the control mode that was best suited was a static control modality.

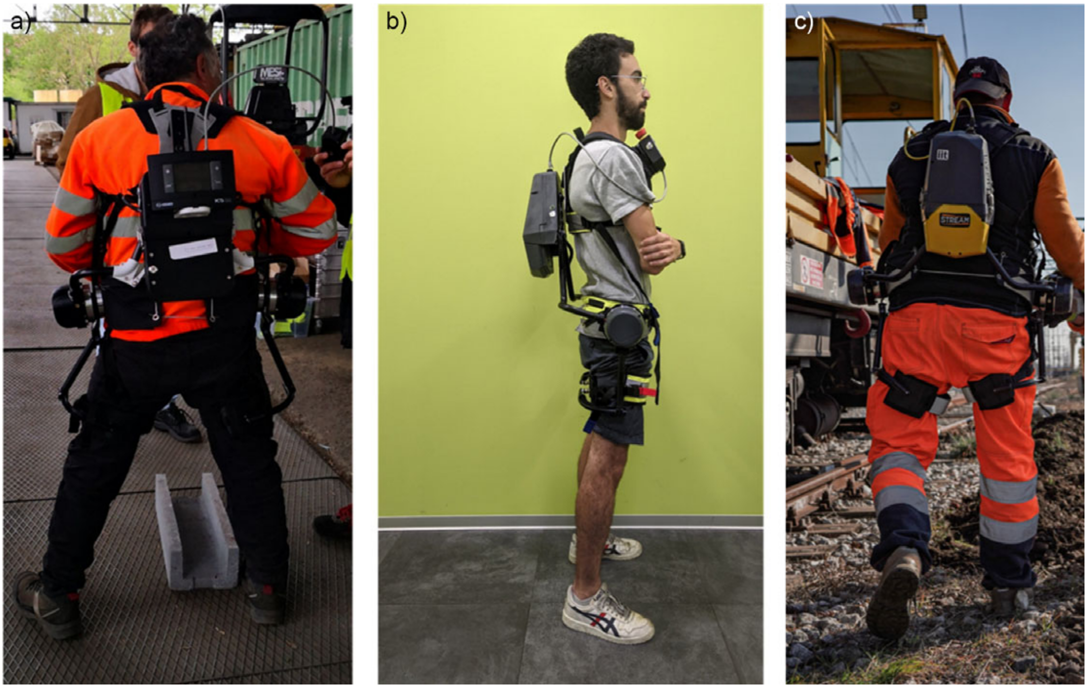


***Figure 9.*** *StreamEXO design evolution with versions 1 (a), 2 (b), and 3 (c) across the process stages.*

### *4.1. Limitations*

This multistage assessment enabled us to gather a significant amount of data, users' feedback, and results; however, there are still areas for future work and development. Although these trials involved nine workers using the system for extended times, more extensive (in terms of test subjects) and longer trials are needed to fully evaluate the user impacts. In future work, we will extend the exoskeleton-wearing period to full (8 hr) shifts, with trials lasting for weeks to months.

A second limitation was the number of workers engaged in the study: 5 for the simulation in a real environment and 4 for the assessment in an Operational environment. Testing with more workers would improve the statistical significance of the results, smoothing some of the trends measured in this work. However, it is also important to note that in the construction sector and the railway maintenance and renewal operations, teams are often composed of a certain number of works (4–6 in this case), and it may be valuable to study impacts across multiple teams, and in different companies that may use different practices to achieve the same end goals. Gender balance in the rail sector is a limitation of the nature of this work. Women (and indeed younger) workers are relatively unusual in the real environment, limiting their inclusion in testing and the generality of any measurement and assessment considerations achieved in device acceptance. Future works will investigate the benefit of exoskeletons in a more gender and population-balanced pool of workers. It should also be noted that one possible suggested benefit of the use of exoskeletons is that it can reduce the "heavy manual work" nature of many tasks opening them up to the more universal population (e.g., sex, age, and size). This is an area of possible exploration for the future.

For the test duration, particularly for Stage 3, where the execution time was measured, the results show that 90 min of the test led to a significant increase in the duration of the execution of each cycle for the NOE modality. When using the EXO modality, the workers show consistent execution times even after 90 min, suggesting lower levels of fatigue and better endurance; however, further extensive testing will be required before this can be definitively proven. This aspect, of course, is also closely related to level productivity, which, we believe, could be strongly modified using the exoskeleton; however, once again, although these results are encouraging and show significant potential, they are currently only preliminary and more detailed and extensive testing is needed in the future.

Finally, developing a sector-specific exoskeleton can limit its applicability outside the selected industrial context. Thus, as is often found, a balance must be struck between the need for generic but less efficient designs versus specific but less applicable custom systems. In all instances, this should be assessed on a case-by-case basis.

## 5. Conclusions

This work, combined with Di Natali et al. (2024a,b), Fanti et al. (2024), and Sposito et al. (2024), shows that a different paradigm that aims to develop a sector-specific instead of a generic industrial exoskeleton could be a valid compromise. On one side, the device can effectively address a limited number of specific industrial needs but, on the other, enhance performance and acceptance. The second contribution of this paper is a comprehensive assessment approach in the format of a multistage experimental campaign to fully characterize the performances and acceptance of a back-support exoskeleton for the rail industry. StreamEXO was tested throughout three experimental stages, including laboratory validation, simulation in the relevant environment, and assessment in an operational environment. In all instances, tests were carried out by real workers. The results combine quantitative and qualitative measurements from nine workers who used the system for multiple days for 80 hr in total.

The work shows that using StreamEXO can reduce worker fatigue by 16.9% for metabolic consumption and by 51% for perceived effort. In parallel, qualitative questionnaires showed that the workers, when using the exoskeleton for this application, had a very high acceptance rate of 88% (10% higher than in comparable studies for the active exoskeleton XoTrunk). Finally, this study suggests that the use of dedicated design may help reduce the challenges and difficulties imposed by the target industrial sector without hindering everyday workflows. These are very encouraging results although future efforts with larger test groups will be required to fully determine the exact nature and extent of the benefits that can at obtained by this approach.

This work provides guidelines for deploying experimental campaigns in industrial sectors to assess the benefit of OEs for workers from TRL5 to TRL7. A further long-term assessment would be required to establish this technology's benefits and acceptance fully, and this will form a focus for future studies.

**Supplementary material.** The supplementary material for this article can be found at http://doi.org/10.1017/wtc.2024.28.

**Data availability statement.** All data can be made available to interested researchers upon request by email to the corresponding author.

**Acknowledgements.** We would like to thank Dr Luca Dellera (ErgoDesign) for contributing to the ergonomic assessment and Laura Masullo (MERMEC STE) for coordinating and organizing the onsite tests. We would also like to thank Rete Ferroviaria Italiana (RFI) for providing access to the railway worksite and the workers of ELEN s.p.a. for their collaboration and willingness to participate.

**Authorship contributions.** Conceptualization: C.D.N.; Funding acquisition: C.D.N.; Investigation: C.D.N., T.P., M.S., V.F., S.L.; Methodology: C.D.N.; Project administration: C.D.N.; Supervision: C.D.N., D.G.C.; Visualization and data curation: C.D.N.; Writing – original draft: C.D.N. Writing – review and editing: D.G.C., C.D.N., T.P., M.S., V.F., S.L.

**Funding statement.** This work was supported by the STREAM project funded by Shift2Rail Joint Undertaking, established under the European Unions Horizon 2020 framework program for research and innovation, under Grant Agreement No 101015418. Responsibility for the information and views expressed in the paper lies entirely with the authors.

**Competing interest.** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. C. Di Natali, T. Poliero, and M. Sposito are the founders of the start-up Proteso, but the StreamEXO system is not a product associated with that start-up.

**Ethical standard.** This work involves human participants and data acquisition from an experimental study. The test was performed in accordance with the experimental protocol approved by the Ethics Committee of Liguria, Italy (protocol number: 001/2019) and complied with the Helsinki Declaration.

In addition, all subjects have read a consent form concerning the test we administered to them, and all of them gave the signed authorization of gathered data, video, and pictures of them during the tests to be used for scientific purposes.

## References


**Ahmad J**, **Fanti V**, **Caldwell DG and Di Natali C** (2024) A framework for the adoption, evaluation and impact of occupational exoskeletons at different technology readiness levels: A systematic review." *Robotics and Autonomous Systems* **179**, 104743.

**Ármannsdóttir AL**, **Beckerle P**, **Moreno JC**, **van Asseldonk EHF**, **Manrique-Sancho M-T**, **Del-Ama AJ**, **Veneman JF and Briem K** (2020) Assessing the involvement of users during development of lower limb wearable robotic exoskeletons: A survey study. *Human Factors* **62**(3), 351–364

**Borg G** (1998a) *The Borg CR10 Scale: A Method for Measuring Intensity of Experience, Eg, Perceived Exertion and Pain*. Stockholm: Stockholm University.

**Borg, G** (1998b) *Borg's Perceived Exertion and Pain Scales*. Champaign, IL: Human Kinetics. https://psycnet.apa.org/record/1998-07179-000.

**Bosch T**, **van Eck J**, **Knitel K and de Looze M** (2016) The effects of a passive exoskeleton on muscle activity, discomfort and endurance time in forward bending work. *Applied Ergonomics* **54**, 212–217

**Cavatorta MP**, **Spada EG and Karam J** (2019) Human-centred approach in Industry 4.0: Virtual validation of the design and application of an exoskeleton in the automotive industry. PhD diss., Politecnico di Torino.

**Colombini D and Occhipinti E** (2020) *ERGOCHECK for a Preliminary Mapping of Risk at Work: Tools, Guidelines, and Applications*. Boca Raton, Florida: CRC Press.

**Crea S**, **Beckerle P**, **De Looze M**, **De Pauw K**, **Grazi L**, **Kermavnar T**, **Masood J**, **O'Sullivan LW**, **Pacifico I and Rodriguez-Guerrero C** (2021) Occupational exoskeletons: A roadmap toward large-scale adoption. Methodology and challenges of bringing exoskeletons to workplaces. *Wearable Technologies* **2**, e11.

**Dawes HN**, **Barker KL**, **Cockburn J**, **Roach N**, **Scott O and Wade D** (2005) Borg's rating of perceived exertion scales: Do the verbal anchors mean the same for different clinical groups? *Archives of Physical Medicine and Rehabilitation* **86**(5), 912–916

**De Bock S**, **Ghillebert J**, **Govaerts R**, **Tassignon B**, **Rodriguez-Guerrero C**, **Crea S**, **Veneman J**, **Geeroms J**, **Meeusen R and De Pauw K** (2022) Benchmarking occupational exoskeletons: An evidence mapping systematic review. *Applied Ergonomics* **98**, 103582.

**De Looze MP**, **Bosch T**, **Krause F**, **Stadler KS and O'Sullivan LW** (2016) Exoskeletons for industrial application and their potential effects on physical work load. *Ergonomics* **59**(5), 671–681.

**DeBlois JP**, **White LE and Barreira TV** (2021) Reliability and validity of the COSMED K5 portable metabolic system during walking. *European Journal of Applied Physiology* **121**, 209–217.

**Di Natali C**, **Buratti G**, **Dellera L and Caldwell D** (2024a) Equivalent weight: Application of the assessment method on real task conducted by railway workers wearing a back support exoskeleton. *Applied Ergonomics* **118**, 104278.

**Di Natali C**, **Chini G**, **Toxiri S**, **Monica L**, **Anastasi S**, **Draicchio F**, **Caldwell DG and Ortiz J** (2021) Equivalent weight: Connecting exoskeleton effectiveness with ergonomic risk during manual material handling. *International Journal of Environmental Research Public Health* **18**(5), 2677.

**Di Natali C**, **Mattila J**, **Kolu A**, **De Vito P**, **Gauttier S**, **Morata M**, **Garcia M and Caldwell D** (2023) Smart tools for railway inspection and maintenance work, performance and safety improvement. *Transportation Research Procedia* **72**, 3070–3077.

**Di Natali C**, **Poliero T**, **Fanti V**, **Sposito M and Caldwell DG** (2024b) Dynamic and static assistive strategies for a tailored occupational Back-support exoskeleton: Assessment on real tasks carried out by railway workers. *Bioengineering* **11**(2), 172.

**Di Natali C**, **Poliero T and Sposito M** (2022) Exoskeleton to support a user's back, edited by Istituto Italiano Di Tecnologia. EU Patent.

**Di Natali C**, **Toxiri S**, **Ioakeimidis S**, **Caldwell DG and Ortiz J** (2020) Systematic framework for performance evaluation of exoskeleton actuators. *J Wearable Technologies* **1**, e4.

**European Commission** (2019) Report from the Commission to the European Parliament and the Council: Sixth Report on Monitoring Development of The Rail Market. *The EU Commission to the European Parliament and the Council*. Brussels, Belgium: European Commission.

**Eva Heckl IP** (2010) *WIR – Women in Rail. Good Practices and Implementation Guide*. Brussels, Belgium: Railway and Infrastructure Companies (CER).

**Fanti V**, **Leggieri S**, **Poliero T**, **Sposito M**, **Caldwell DG and Di Natali C** (2024) Multi-exoskeleton performance evaluation: Integrated muscle energy indices to determine the quality and quantity of assistance." *Bioengineering* **11**(12), 1231.

**Fox RR**, **Lu M-L**, **Occhipinti E and Jaeger M** (2019) Understanding outcome metrics of the revised NIOSH lifting equation. *Applied Ergonomics* **81**, 102897.

**Gillette JC**, **Saadat S and Butler T** (2022) Electromyography-based fatigue assessment of an upper body exoskeleton during automotive assembly. *Wearable Technologies* **3**, e23.

**Grazi L**, **Chen B**, **Lanotte F**, **Vitiello N and Crea S** (2019) Towards methodology and metrics for assessing lumbar exoskeletons in industrial applications. In *2019 II Workshop on Metrology for Industry 4.0 and IoT (MetroInd4. 0&IoT)*.

**Héder M** (2017) From NASA to EU: The evolution of the TRL scale in public sector Innovation. *The Innovation Journal* **22**(2), 1–23.

**Hellig T**, **Rick V**, **Mertens A**, **Nitsch V and Brandl C** (2019) Investigation of observational methods assessing workload of static working postures based on surface electromyography. *Work* **62**(2), 185–195.

**Kermavnar T**, **de Vries AW**, **de Looze MP and O'Sullivan LW** (2021) Effects of industrial back-support exoskeletons on body loading and user experience: An updated systematic review. *Ergonomics* **64**(6), 685–711.

**Kim YG**, **Little K**, **Noronha B**, **Xiloyannis M**, **Masia L and Accoto D** (2020) A voice activated bi-articular exosuit for upper limb assistance during lifting tasks. *Robotics and Computer-Integrated Manufacturing* **66**, 101995.

**Kim S**, **Nussbaum MA and Smets M** (2022) Usability, user acceptance, and health outcomes of arm-support exoskeleton use in automotive assembly: An 18-month field study. *Journal of Occupational and Environmental Medicine* **64**(3), 202–211.

**Kim S**, **Nussbaum MA**, **Smets M and Ranganathan S** (2021) Effects of an arm-support exoskeleton on perceived work intensity and musculoskeletal discomfort: An 18-month field study in automotive assembly. *American Journal of Industrial Medicine* **64**(11), 905–914.

**Kopp V**, **Holl M**, **Schalk M**, **Daub U**, **Bances E**, **García B**, **Schalk I**, **Siegert J and Schneider U** (2022) Exoworkathlon: A prospective study approach for the evaluation of industrial exoskeletons. *Wearable Technologies* **3**, e22.

**Kuber PM**, **Abdollahi M**, **Alemi MM and Rashedi E** (2022) A systematic review on evaluation strategies for field assessment of upper-body industrial exoskeletons: Current practices and future trends. *Annals of Biomedical Engineering Rashedi* **50**(10), 1203–1231.

**Marras WS** (2006) *Fundamentals and Assessment Tools for Occupational Ergonomics*. Boca Raton, Florida: CRC Press.

**McFarland T and Fischer S** (2019) Considerations for industrial use: A systematic review of the impact of active and passive upper limb exoskeletons on physical exposures. *IISE Transactions on Occupational Ergonomics and Human Factors* **7**(3–4), 322–347.

**Monk A** (2000) User-centred design: The home use challenge. In Sloane A and van Rijn F (eds), *Home Informatics and Telematics: Information, Technology Society*. Boston: Kluwer Academic Publishers, pp. 181–190.

**NIOSH** (1981) Work practice guide for manual load lifting. NIOSH Technical Report No. 81-122. US National Institute for Occupational Safety and Health Cincinnati.

**O'Sullivan LW**, **Power V**, **De Eyto A and Ortiz J** (2017) User centered design and usability of bionic devices. In *Converging Clinical and Engineering Research on Neurorehabilitation II: Proceedings of the 3rd International Conference on NeuroRehabilitation (ICNR2016)*, October 18–21, 2016, Segovia, Spain.

**Occhipinti E** (1998) OCRA: A concise index for the assessment of exposure to repetitive movements of the upper limbs. *Ergonomics* **41**(9), 1290–1311.

**Onofrejová D and Balážiková M** (2021) Simulation model for evaluation of ergonomic load in the use of exoskeletons. *IOP Conference Series Materials Science and Engineering* **1199**(1), 012011.

**Otten BM**, **Weidner R and Argubi-Wollesen A** (2018) Evaluation of a novel active exoskeleton for tasks at or above head level. *IEEE Robotics and Automation Letters* **3**(3), 2408–2415.

**Parasuraman A** (2000) Technology readiness index (TRI): A multiple-item scale to measure readiness to embrace new technologies. *Journal of Service Research* **2**(4), 307–320.

**Pesenti M**, **Antonietti A**, **Gandolla M and Pedrocchi A** (2021) Towards a functional performance validation standard for industrial low-back exoskeletons: State of the art review. *Sensors* **21**(3), 808.

**Schneider E**, **Copsey S and Irastorza X** (2010) *OSH [Occupational Safety and Health] in Figures: Work-Related Musculoskeletal Disorders in the EU-Facts and Figures*. Luxembourg: Office for Official Publications of the European Communities.

**Spada S**, **Ghibaudo L**, **Gilotta S**, **Gastaldi L and Cavatorta MP** (2018) Analysis of exoskeleton introduction in industrial reality: Main issues and EAWS risk assessment. In *Advances in Physical Ergonomics and Human Factors: Proceedings of the AHFE 2017 International Conference on Physical Ergonomics and Human Factors*, July 17–21, 2017, the Westin Bonaventure Hotel, Los Angeles, California, USA, 8.

**Sposito M**, **Di Natali C**, **Toxiri S**, **Caldwell DG**, **De Momi E and Ortiz J** (2020) Exoskeleton kinematic design robustness: An assessment method to account for human variability. *Wearable Technologies*, **1**, e7.

**Sposito M**, **Fanti V**, **Poliero T**, **Caldwell DG and Di Natali C** (2024) Field assessment of active BSE: Trends over test days of subjective indicators and self-reported fatigue for railway construction workers. *Heliyon* **10**(12), e33055.

**Sposito M**, **Poliero T**, **Di Natali C**, **Toxiri S**, **Anastasi S**, **Draicchio F**, **Monica L**, **Caldwell D**, **De Momi E and Ortiz J** (2021) Physical comfort of occupational exoskeletons: Influence of static fit on subjective scores. In: *The 14th PErvasive Technologies Related to Assistive Environments Conference*.

**Sylla N**, **Bonnet V**, **Colledani F and Fraisse P** (2014) Ergonomic contribution of ABLE exoskeleton in automotive industry. *International Journal of Industrial Ergonomics* **44**(4), 475–481.

**Takala E-P**, **Pehkonen I**, **Forsman M**, **Hansson G-Å**, **Mathiassen SE**, **Neumann WP**, **Sjøgaard G**, **Veiersted KB**, **Westgaard RH and Winkel J** (2010) Systematic evaluation of observational methods assessing biomechanical exposures at work. *Scandinavian Journal of Work Environment and Health* **36**, 3–24.

**Theurel J and Desbrosses K** (2019) Occupational exoskeletons: Overview of their benefits and limitations in preventing work-related musculoskeletal disorders. *IISE Transactions on Occupational Ergonomics and Human Factors* **7**(3–4), 264–280.

**Torricelli D and Pons JL** (2019) Eurobench: Preparing robots for the real world. In *Wearable Robotics: Challenges and Trends: Proceedings of the 4th International Symposium on Wearable Robotics, WeRob2018*, October 16–20, 2018, Pisa, Italy, 3.

**Torricelli D**, **Rodriguez-Guerrero C**, **Veneman JF**, **Crea S**, **Briem K**, **Lenggenhager B and Beckerle P** (2020) Benchmarking wearable robots: Challenges and recommendations from functional, user experience, and methodological perspectives. *Frontiers in Robotics and AI* **7**, 561774.

**Tsao L and Ma L and Papp C-T** (2018) Using non-invasive wearable sensors to estimate perceived fatigue level in manual material handling tasks. In *International Conference on Applied Human Factors and Ergonomics*. Cham: Springer, pp. 65–74.
**Van der Grinten MP** (1991) Test-retest reliability of a practical method for measuring body part discomfort. *Designing for Everyone* **1**, 54–57.
**van der Have A**, **Wang W**, **Van Rossom S and Jonkers I** (2023) Joint loading topography during occupational tasks – a musculoskeletal modeling approach to substantiate ergonomic recommendations. *International Journal of Industrial Ergonomics* **95**, 103451.
**Waters TR**, **Lu M-L**, **Piacitelli LA**, **Werren D and Deddens JA** (2011) Efficacy of the revised NIOSH lifting equation to predict risk of low back pain due to manual lifting: Expanded cross-sectional analysis. *Journal of Occupational and Environmental Medicine* **53**(9), 1061–1067.
**Waters T**, **Occhipinti E**, **Colombini D**, **Alvarez-Casado E and Fox R** (2016) Variable lifting index (VLI): A new method for evaluating variable lifting tasks. *Human Factors* **58**(5), 695–711.
**Waters TR**, **Putz-Anderson V and Garg A** (1994) *Applications Manual for the Revised NIOSH Lifting Equation*. Cincinnati, OH: Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health.
**Waters TR**, **Putz-Anderson V**, **Garg A and Fine LJ** (1993) Revised NIOSH equation for the design and evaluation of manual lifting tasks. *Ergonomics* **36**(7), 749–776.
**Zelik KE**, **Nurse CA**, **Schall Jr MC**, **Sesek RF**, **Marino MC and Gallagher S** (2022) An ergonomic assessment tool for evaluating the effect of back exoskeletons on injury risk. *Applied Ergonomics* **99**, 103619.

**Cite this article:** Di Natali C, Poliero T, Sposito M, Fanti V, Leggieri S and Caldwell DG (2025) From the idea to the user: a pragmatic multifaceted approach to testing occupational exoskeletons. *Wearable Technologies*, 6, e5. doi:https://doi.org/10.1017/wtc.2024.28