



A VR-based approach in conducting MTM for manual workplaces

Valentina Gorobets¹ · Valentin Holzwarth² · Christian Hirt¹ · Noëlle Jufer³ · Andreas Kunz¹

Received: 11 December 2020 / Accepted: 7 May 2021 / Published online: 29 June 2021
© The Author(s) 2021

Abstract

Due to current trends in the manufacturing industry, such as mass customization, manual operations contribute drastically to the overall costs of a product. Methods-Time-Measurement (MTM) identifies the optimization potential of manual workplaces, which significantly influences a worker's productivity. However, traditional MTM requires great efforts to observe and transcribe manual assembly processes. Yet, various digital approaches exist that facilitate MTM analyses. While most of these approaches require the existence of real workplaces or cardboard mock-ups, it would be beneficial to conduct a virtual MTM in earlier phases of production planning. However, the quality of virtual MTM analyses compared to traditional MTM conducted in reality has not been assessed yet. This paper is addressing it by conducting a comparative user study with 21 participants completing the same task both at a real and virtual workplace, which they access via virtual reality technology. Our results show that participants' MTM-2 values achieved at the VR workplace are comparable to those at the real workplace. However, time study data reveals that participants moved considerably slower in VR and thus needed more time to accomplish the task. Consequently, for the measurement of manual work in VR, it is even necessary to utilize predetermined times, such as MTM-2 since time study data is insufficient. This paper also serves as a proof of concept for future studies, investigating automated transcription systems that would further decrease the efforts conducting MTM analyses.

Keywords Workplace optimization · MTM · Human-centered design · Virtual reality

1 Introduction

Manual operations in a production substantially contribute to the overall costs of a product and should therefore be

measured and evaluated carefully [6]. This becomes increasingly challenging as products change more rapidly and workplaces are decentralized. Thus, an evaluation on the machine level only (see Fig. 1) is not sufficient anymore. Recently, worker behavior cannot be assumed to be static, as it also includes locomotion, mainly driven by new production strategies such as Chacu-Chacu or decentralized production [18]. Thus, also cell level, shop floor level, and even factory level should be considered when evaluating manual work efficiency [1].

Figure 1 shows a typical structure of a manufacturing company. From top to bottom, the level of detail increases regarding manual work that can be analyzed. Often, the details in the description of manual work on the enterprise and factory level are not sufficient, which makes it difficult to conduct an optimization. Therefore, manual worker operations are mainly evaluated on machine level, at fixed workplaces, by applying so-called MTM (Methods-Time Measurement). MTM is one of the most widely used predetermined time measurement systems, which defines a particular time measured in so-called Time Measurement Units (TMUs) for each basic operation described in it [6].

✉ Valentina Gorobets
gorobets@iwf.mavt.ethz.ch

Valentin Holzwarth
valentin.holzwarth@rhysearch.ch

Christian Hirt
hirtc@iwf.mavt.ethz.ch

Noëlle Jufer
jufer@acel.ch

Andreas Kunz
kunz@iwf.mavt.ethz.ch

¹ ETH Zurich, Leonhardstrasse 21, 8092 Zurich, Switzerland

² RhySearch, Werdenbergstrasse 4, 9471 Buchs, Switzerland

³ Acél & Partner, Aargauerstrasse 1, 8048 Zurich, Switzerland

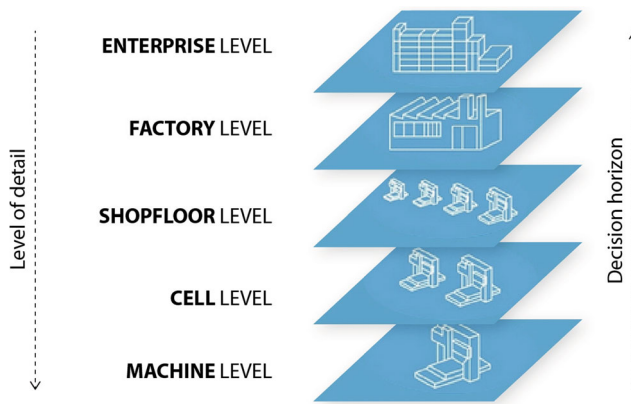


Fig. 1 Various investigation levels for manual work, adapted from [27]

One TMU is equivalent to 0.036 s, which allows for an in-depth analysis even for basic operations in the sub-second range. MTM has different standards, and each standard has its own set of basic operations and predetermined times.¹

The most comprehensive MTM standard is MTM-1 which consists of the following basic operations: reach, grasp, release, move, position, apply pressure, disengage, turn, crank, visual inspection, and body, leg, and foot movements. MTM-1 is not only complex but also precise due to the high level of detail. This is particularly favored in cases of mass production that should be analyzed with the highest possible level of detail. If such a detailed work description is not required, e.g., for a smaller production pipeline, the MTM-2 standard is used. MTM-2 is based on MTM-1 and consists of the following basic operations: get, put, apply pressure, regrasp, crank, eye action, foot motion, step, and bend and arise. MTM-2 is less precise than MTM-1, but it also allows conducting faster MTM analyses. For even smaller production entities, e.g., for batch production, the so-called MTM-MEK is used. It consists of the following basic operations: get and place, place, handle tool, operate, motion cycles, body motion, and visual control. And finally, to make MTM more universally applicable, the MTM-UAS standard was introduced. It has the same basic operations as MTM-MEK but with even less detailed information for each basic motion. A brief overview of the abovementioned MTM standards is presented in Fig. 2.

When analyzing manual operations with MTM, a video of a work process is recorded and manually transcribed. This process is time-consuming due to the need for the subsequent manual video transcription, which has to be conducted by an MTM expert. Furthermore, the MTM analysis is frequently done when the workplace is already built and can hardly be altered anymore. To overcome this issue, cardboard mock-ups are used in the planning phase,

¹<http://mtm-international.org/technical-platform/>

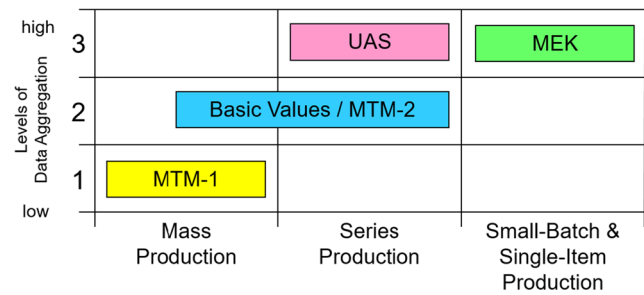


Fig. 2 Application fields for different MTM standards based on the IMD Technical Platform

which represent the workplace to be investigated in full-size. Besides the time, cost, and generated waste, cardboard engineering often lacks details, which makes it difficult for workers to perform in the same way as they would do at a real workplace. Such details include raw materials and intermediates, which are usually not available during planning, to be processed or assembled by a worker.

However, novel opportunities arise due to assembly lines and production plants increasingly being planned and designed virtually. This is based on virtual planning data, which is usually available in the form of three-dimensional geometry (e.g., computer-aided design data), could be harnessed to generate a virtual environment (VE), wherein workers perform the manual operations to be analyzed virtually. This can be achieved through virtual reality (VR) technology, which immerses users within a VE, while allowing them to interact with virtual objects through the utilization of the controllers. Few prior works investigated the potentials of VR technology to support manual work measurement with a particular focus on MTM analysis. However, the setups in prior works are characterized by numerous limitations, such as not allowing for real walking, non-intuitive interaction, and the weightlessness of virtual objects [10, 11]. Consequently, the real work procedure cannot be fully replicated, and the results of work measurement in VR might be only of limited benefit [5].

Within this paper, we address this by developing a VE that emulates a typical workplace for manual assemblies, in which a worker can naturally walk, while completing a given task. The goal of the paper is to investigate work measurement in VR, including an MTM analysis and time study, compared to work measurement at the corresponding physical workplace.

To achieve this, we conduct a comparative study, wherein we analyze 21 study participants performing the same task in a VE and at a physical workplace using an MTM-2 approach and time study. We organize the remainder of this work as follows: Section 2 introduces the related work, Section 3 describes the task and the technical setup. The results are described in Section 4. Section 5 discusses the

achieved results and, finally, Section 6 concludes the paper, and an outlook is given in Section 7.

2 Related work

The measurement of manual work in the manufacturing industry is associated with various benefits, including productivity gains and quality improvements [6]. With growing mass customization in an automated or semi-automated work environment, manual assembly tasks also grow in importance. Depending on a product's configuration, manufacturing and assembly tasks can vary in time, and thus current research focuses on decentralized production.

In such decentralized production environments, so-called “slow” products do not block “faster” ones anymore, and thus the overall output of production can be kept stable or even increased, although multiple product variants are produced. However, decentralized production needs not only a careful planning but also a thorough investigation of manual assembly tasks, including walking and human-robot interaction. Moreover, there is no such support for planning manual work; and thus, MTM is still based on traditional cardboard engineering.

Since most of the 3D models for machines and tools already exist, it seems reasonable to employ VR technology to conduct MTM analyses. An early approach to digitally evaluate an ergonomic design was introduced by Schmidt and Wendt [25]. The software tool “COSIMAN” (Computer Simulation of Manual Assembly) was able to evaluate basic designs but was limited to more complex scenarios. Moreover, the handling of the software tool was too complex for industrial applications. Further applications of COSIMAN were also introduced by Kummelsteiner [16]. Another VR-supported MTM approach was introduced by Chan [7], who simulated a manual assembly line and retrieved MTM values by using virtual humans (avatars). However, the control of the avatars by a computer mouse, as well as missing functionalities for the presentation of products and process designs, imposed a substantial effort to achieve simulation results.

In spite of its drawbacks regarding operability, there exists a wide range of digital human model programs, as stated by Mühlstedt et al. [21]; they all offer an exocentric view on the manufacturing task. The main focus of these virtual human simulation programs is to analyze human postures and to determine the workplace, as described by Yang et al. [30]. Furthermore, these programs should assess the visibility and accessibility of an operation as stated by Chedmail et al. [8] or evaluate postures as stated by Bubb et al. [4]. In these virtual human programs, conventional MTM also can be integrated, as well as posture analysis techniques [13].

Based on inverse kinematics, the physical strain on each joint can be calculated for any given operational and external load [29].

Since motion capture systems became available, motion tracking methods for tracking an operator's real movements to control the manikin became popular due to its realistic feeling and outstanding, lifelike realism. Already, in 2000, Chryssolouris et al. [9] proposed a “virtual assembly work cell,” which allows natural interaction with the VE to perform an assembly task. However, the overall system complexity and the particular limitations of the tracking system only allowed for a limited range of applications. Instead of a consecutive definition of the avatar's movement, newer systems allow for a direct coupling of the avatar to the workers' movements. For doing so, the user wears components of an optical tracking system, a data glove, and a data helmet.

In 2005, Jianping and Keyi [15] tracked the real operator's movement in real time to control the virtual human “JACK” within a virtual maintenance system. Furthermore, a virtual assembly design environment was developed by Jayaram et al. [14], who also uses a virtual human that can be directly controlled by the worker's movements. The motion data of the real operator was recorded and imported into JACK to perform an ergonomic analysis. Later, Wu et al. [28] used data gloves and a field-of-view tracker to capture the operator's hand for a virtual assembly path planning in a VE. In 2012, Osterlund and Lawrence [23] used full-body optical motion tracking to control avatars in an astronaut training. However, their setup was dependent on a well-defined physical setup, much like in the abovementioned cardboard installation. According to various sources [19, 26], establishing a VE with complete physical attributes is still difficult and time-consuming, since such systems require a long setup time, in particular for optical, outside-in tracking systems. Following Seth et al. [26], also a simulation of realistic interaction using haptic devices is still difficult for the virtual prototyping community, while sensations, such as real walking to perceive distances, are completely missing. Hence, it is still inconvenient for a real operator to control a virtual human. Following Qiu et al. [24], current human motion capturing technology substantially limits the range of its applications.

With the spread of information technology, different approaches for analyzing workers' performance and training new workers had appeared. Benter et al. [3] introduced an approach that used a 3D camera to capture data and analyzed working time with the MTM-1 method. The workplace consisted of three workstations, and the worker was assembling gearboxes for the duration of 20 min. Ma et al. [20] came up with a framework to evaluate manual work operations with the support of motion tracking. For these

purposes, a marker-based optical tracking system with a total of 13 markers was used. The system consisted of a head-mounted display (HMD), a data glove, and eight cameras for body motion and hand-gesture motion recognition. To evaluate workers' performance, a so-called Maynard operation sequence technique (MOST) was planned for use. There are three motion groups in the MOST system: general move, controlled move, and tool use. To validate the technical feasibility of the proposed framework, only two work tasks were taken: lifting an object and pushing a button. No walking was included. Another approach to automatically monitor the execution of human-based assembly operations was proposed by Andrianakos et al. [1]. Their approach is based on machine learning techniques and utilizes a Single Shot Detector algorithm for object detection. This algorithm detects objects of multiple categories and sizes in real time. For data collection, a single vision sensor (i.e., a webcam) was used. To evaluate this approach, a simple three-step assembly task was proposed. An important restriction of this method is that a task must be completed in a defined specific order. Müller et al. [22] designed a Smart-Assembly-Workplace (SAW) which was used in a bicycle e-hub assembly. SAW was designed to share knowledge about the assembly sequence with less qualified workers in an intuitive way. SAW consists of a combined working and learning environment. To define the learning sequence and time, MTM is used. For data retrieval, the motion-sensing device Microsoft Kinect is installed on top of the workplace, facing downwards, which allows the determination of the worker's hand's position without using markers. However, the reliability of Kinect depends on illumination conditions. It was reported that the hand tracking often fails to correctly locate the hands if illumination conditions vary. A tree-based approach to recognize MTM-UAS codes in VR was proposed by Bellarbi et al. [2]. It captures the tracking data of an HMD and controllers and divides this data into small sequences of movements. In this algorithm, all possible body motions belong to one of these three categories: eye movement, body movement, or hand movement. Each small sequence of movements from the captured data is compared to the data from the algorithm tree in order to get a corresponding to this sequence MTM-UAS code.

First approaches by Kunz et al. [17] showed that also real walking can be integrated in such virtual interpretations. However, a comparison to the real-world counterpart is still missing. However, the quality of MTM analyses and time study conducted in VR compared to the analysis of a corresponding real world task should be investigated [5, 11]. This paper will thus contribute to the topic of performing MTM analyses completely in VR by comparing the findings to a similar real-world workplace.

3 Methodology

In this section, we describe the methodology of this paper, including a user study, where each participant performed the same task in VR and reality. Furthermore, we outline the data retrieval and analysis procedure to conduct the intended comparison of work measurement in reality and VR.

3.1 Participants

The user study consists of 21 participants: 6 female and 15 male, recruited from the university's student body. Since MTM is designed for a workplace evaluation by experienced workers, we had to design the task for the user study suitable for participants without any prior training. As the task was consciously chosen to be very simple, any biasing effect by learning is unlikely. Thus, each participant performed the same task twice, once at the real workplace and once in VR. To avoid any biasing from the task sequence, participants either started with the real task or with the virtual task in alternating order.

3.2 Technical setup

The technical setup consists of two identical workplaces in reality (see Fig. 4 left) and VR (see Fig. 4 right). Participants access the VR workplace with the HTC Vive Pro system. It uses so-called lighthouses to track users' head position and orientation, as well as their hands holding the HTC Vive controllers. In our virtual setup, the participants are using only one controller to manipulate the objects. The controller has additional buttons and a touchpad to allow further interaction with the VE. Pulling the controller's trigger, for instance, performs a grasping action to grab the virtual object with the virtual hand. Users see their virtual hand and the complete VE through the HMD, which also visually disconnects them from the real world. To reach all relevant objects in the VE, users are able to freely walk within a 5m × 5m tracking space. In both the real and the virtual scenery, the users' movements are video-captured for a later manual transcription.

3.3 Task description

To compare the MTM analysis in reality and VR, a task that could be completed by inexperienced study participants was chosen. It is a simplified version of an industrial task, containing basic operations of the MTM-2 standard such as get, put, eye action, step, and bend and arise. In contradiction to MTM-UAS and MTM-MEK, which evaluate the walking time based on geometric distances only, MTM-2 allows counting steps of the user and gives

a predetermined TMU value for each step. This allows evaluating walking behavior based on human behavior and not only on geometrically obtained values.

The top-down layout of the workplace is shown in Fig. 3. At the beginning of each study session, a pre-recorded video in VR with text instructions is shown to the participants. In this video, the task and the sequence of the process steps are introduced. The virtual workplace is designed to fit into the tracking space of $5\text{ m} \times 5\text{ m}$ that is supported by the HTC Vive system. The participants start in front of a palette in 1.8 m distance to the larger table. They grab the box from the floor and put it on the bigger table. Afterward, they grab the screwdrivers one by one from the smaller table and put them into the box. Then they walk to the box lid, grab it, and close the box with the lid. The last step is to take the closed box, walk to the palette, and put this box on the palette. There was no time limit for this task. Participants are asked to complete the task at a natural speed. The real and virtual environments are shown in Fig. 4.

3.4 Data retrieval

Each study participant was recorded while performing the task in reality and in VR. The recording of the real task was conducted through a video camera, while “screen recordings” (i.e., a recording of a participant’s 1st person view) were created for the task in VR. Subsequently, each participant’s movements were manually transcribed according to the corresponding MTM-2 codes. The transcription was done by two people independently. Additionally, a time study was conducted for each user by measuring their task completion time in VR and reality.

4 Results

This section describes the findings of our comparative user study, including MTM-2 analysis and time study for the task conducted in reality and VR.

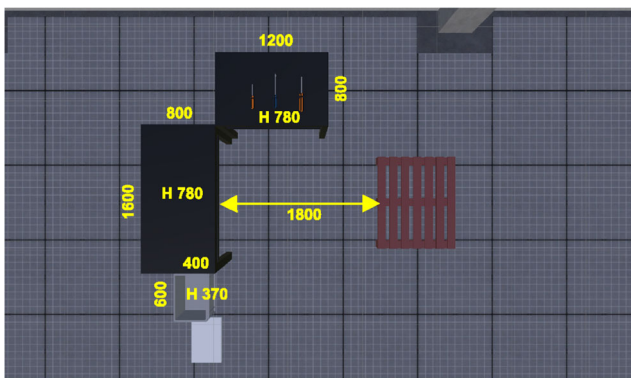


Fig. 3 A top-down layout of the real and virtual workplace, enhanced with the $5\text{ m} \times 5\text{ m}$ VR tracking space

4.1 MTM-2 and time study in reality

Since MTM relies on statistically retrieved values (the “Measured Times”), this first data evaluation is to assess whether the participants of the study performed the given task at a normal working speed or not. For this, we compare the time study and MTM-2 transcription of the task conducted in reality by each participant. The results of this analysis are shown in Table 1 on the right and plotted in Fig. 5.

When analyzing the results of the time study, we see a mean value of $m_{\text{time study, real}} = 27.667\text{ s}$ with a standard deviation of $SD_{\text{time study, real}} = 4.4\text{ s}$. This, in fact, shows that there are considerable differences between each individual study participant. However, when analyzing MTM-2 values of the transcription instead, we achieve $m_{\text{MTM-2, real}} = 27.710\text{ s}$ and $SD_{\text{MTM-2, real}} = 2.4\text{ s}$. This is an expected outcome, since the MTM-2 analysis decouples workers’ movement from their individual speed in performing a task.

4.2 MTM-2 and time study in VR

Similar to the analysis in reality, we analyze both the MTM-2 and time study for each participant performing the task in VR, which is shown in Table 1 on the left and plotted in Fig. 6.

When analyzing the time study from the VR task only, we see a mean value of $m_{\text{time study, vr}} = 36.286\text{ s}$ with a standard deviation of $SD_{\text{time study, vr}} = 7.9\text{ s}$. This shows that the performance between participants substantially differs in VR. Seven users had an MTM-2 time that is larger than the time study, meaning that they performed the actions in VR faster than it is foreseen by the MTM-2 (see Table 1 on the left).

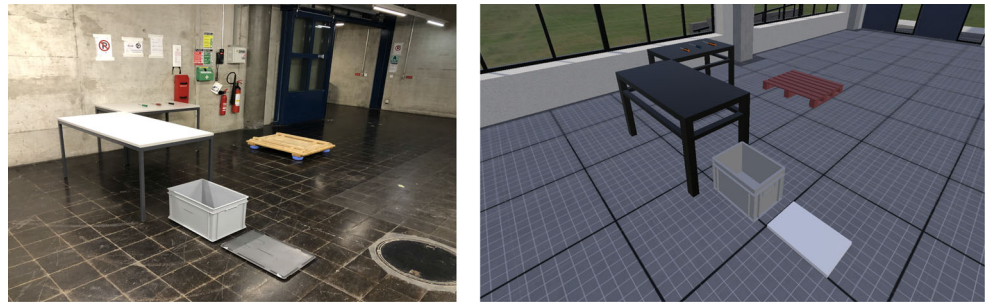
In comparison to the time study, the transcribed MTM-2 values of exactly the same work process are considerably smaller. The MTM-2 mean value for VR is $m_{\text{MTM-2, vr}} = 28.991\text{ s}$, while the standard deviation is $SD_{\text{MTM-2, vr}} = 2.7\text{ s}$. Similarly to the analysis in reality, this is an expected result since MTM-2 decouples the speed of a worker’s movement from the descriptive class of their movement.

4.3 Comparison of results in reality and VR

As it is one strength of MTM-2 to make workplaces comparable, we chose the overall MTM-2 values for the task conducted in reality and VR for the comparison. It is hypothesized that equal MTM-2 values will show that a VR workplace is capable of adequately representing a real workplace.

The overall information about the comparison of the virtual and real task completion time is shown in Table 2. Figure 7 shows that the MTM-2 values in reality and VR are very similar for each user, which confirms our

Fig. 4 Comparison of real and virtual tasks during the user study. Left: the real workplace; right: the virtual counterpart



previously stated hypothesis that the MTM-2 analysis for a task conducted in VR has comparable quality and can, in fact, replace a real one.

However, it is also visible that the values are not absolutely coherent. To further clarify the resulting error, the quotient between the corresponding MTM values for an individual user is calculated, which is referred to as VR MTM-2 accuracy and should ideally be equal to 1. The deviation of the VR MTM-2 accuracy from 1 gives a measure of the quality of the MTM-2 analysis in VR. Figure 8 shows the deviation from this ideal value. For the VR MTM-2 accuracy, we achieve a mean of $m_{VR\ MTM-2\ accuracy} = 1.05$ and $SD_{VR\ MTM-2\ accuracy} = 0.08$.

5 Discussion

While there is already a good correlation between the MTM-2 analysis in reality and VR, we made some observations during the course of the study, which could be an explanation for the typically longer task completion time when being exposed to the VE.

Missing haptic feedback for grasping objects When grasping an object, humans rely on vision for coarsely approaching the object to be grasped, while the decision whether an object is grasped relies on haptic cues only. Such haptic feedback was missing in the VE, and thus users had to

Table 1 Overall completion time for each user measured in VR and reality by direct observation time study and MTM-2

User	VR (MTM-2)		VR (time study) (s)	Reality (MTM-2)		Reality (time study) (s)
	(TMU)	(s)		(TMU)	(s)	
# 1	765	27.518	40	868	31.223	32
# 2	1138	40.935	59	917	32.986	35
# 3	806	28.993	37	776	27.914	25
# 4	702	25.252	27	714	25.683	22
# 5	770	27.698	27	768	27.626	33
# 6	708	25.468	32	698	25.108	26
# 7	829	29.820	34	756	27.194	27
# 8	808	29.065	28	783	28.165	27
# 9	844	30.360	38	808	29.065	27
# 10	804	28.921	41	792	28.800	31
# 11	812	29.209	28	714	25.683	19
# 12	795	28.597	39	670	24.101	20
# 13	755	27.158	32	714	25.683	22
# 14	727	26.151	28	738	26.547	26
# 15	871	31.331	43	801	28.813	36
# 16	804	28.921	41	809	29.101	30
# 17	753	27.086	48	821	29.532	29
# 18	757	27.230	26	771	27.734	27
# 19	791	28.453	38	765	27.518	31
# 20	813	29.245	37	783	28.165	29
# 21	873	31.403	39	711	25.576	27
Mean	806	28.991	36.286	770	27.710	27.667

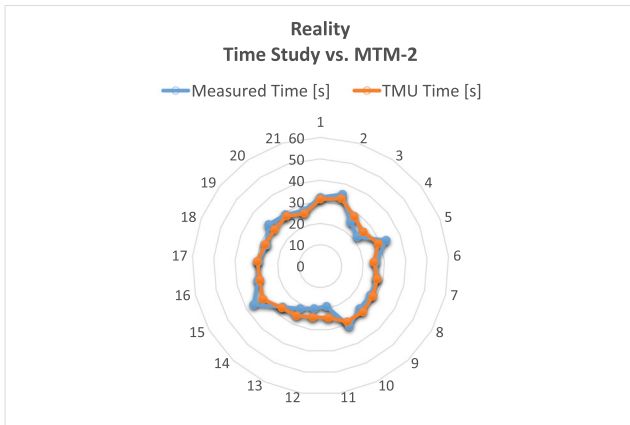


Fig. 5 Comparison of the time study and the transcribed MTM-2 times in reality, shown in seconds

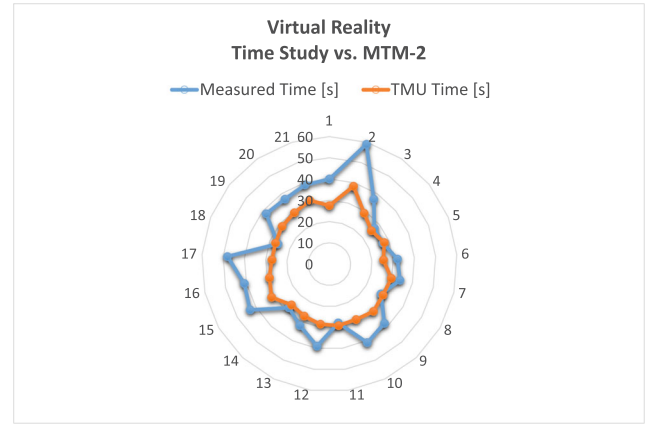


Fig. 6 Comparison of time study and the transcribed MTM-2 times in VR, shown in seconds

observe the object for a yellow wireframe that indicates that the virtual hand touches the virtual object. Although this is a common way in VR to indicate whether objects are touched or not, it imposes an additional cognitive effort, which was

taken into account during the MTM-2 transcription as an “eye action” ($E = 7$ TMUs). However, we noticed that users significantly slow down their physical movement, in particular when trying to grasp the lid from the ground. This is

Table 2 Overall relations of the virtual completion time to the real for each user

User	VR MTM-2 Accuracy	MTM-2 Difference (VR - Reality)		Time Study Difference (VR - Reality)
		[TMU]	[s]	
# 1	0.88	-103	-3.71	8
# 2	1.24	221	7.96	24
# 3	1.04	30	1.08	12
# 4	0.98	-12	-0.43	5
# 5	1.00	2	0.07	-6
# 6	1.01	10	0.36	6
# 7	1.10	73	2.63	7
# 8	1.03	25	0.90	1
# 9	1.04	36	1.30	11
# 10	1.02	12	0.43	10
# 11	1.14	98	3.53	9
# 12	1.19	125	4.50	19
# 13	1.06	41	1.48	10
# 14	0.99	-11	-0.40	2
# 15	1.09	70	2.52	7
# 16	0.99	-5	-0.18	11
# 17	0.92	-68	-2.45	19
# 18	0.98	-14	-0.50	-1
# 19	1.03	26	0.94	7
# 20	1.04	30	1.08	8
# 21	1.23	162	5.83	12
Mean	1.05	32.62	1.28	8.62

The accuracy of the virtual MTM is obtained by dividing the corresponding virtual to the real TMU value

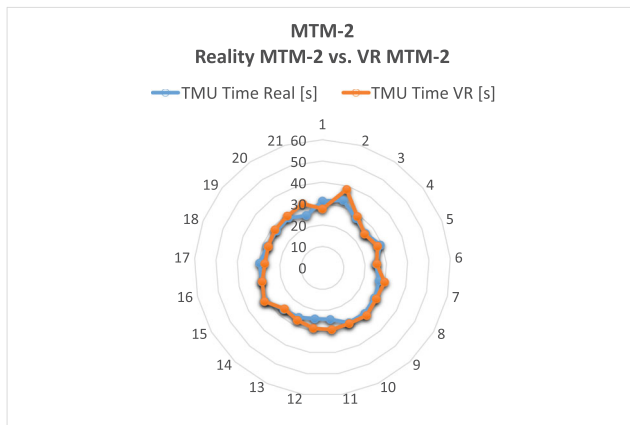


Fig. 7 Comparison of MTM-2 times in VR and reality, shown in seconds

probably due to the fact that users are afraid to collide the hand-held controller with the physical floor.

Missing haptic feedback for placing objects When placing real objects on top of another, humans again rely on haptic feedback, indicating that the two physical objects collided with each other. This then indicates that the grasp can be released to place the object. Because of the VE lacking this feedback channel, we noticed a slowdown of the user's movement in the VE in the following three situations: (i) placing the virtual box on the virtual table, (ii) placing the virtual lid on the virtual box, and (iii) placing the virtual box on the virtual palette.

Unfamiliar navigation means We noticed that participants who never had any VR experience before did not walk naturally when navigating the first distance from the starting position to the box. They either made smaller steps, or generally moved slower without shortening the step size. However, we also noticed that users quickly familiarize with real walking in VR, so that already after this first very short path their walking speed was close to natural again.

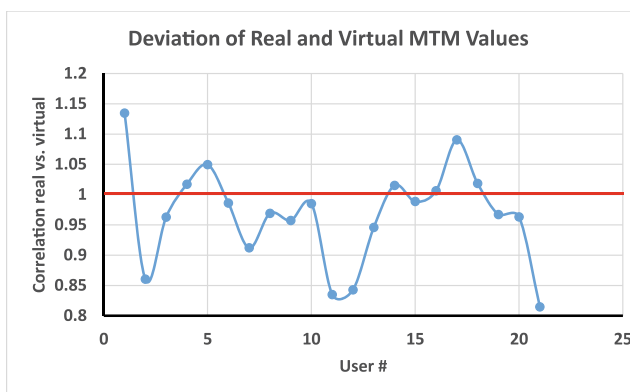


Fig. 8 Deviation of MTM-2 times in VR and reality

6 Conclusion

In this paper, we followed the research question, how well an MTM-2 analysis and a time study could be conducted in VR compared to traditional analyses in reality. Conducting a study in VR brings the benefit of avoiding non-optimal workplace designs already in the planning phase without requiring the construction of physical mock-ups (cardboard engineering) for data acquisition.

Our comparative user study includes two identical workplaces in reality and VR and shows that it is possible to achieve comparable results in both setups using the MTM-2 evaluation. However, the overall completion time that is measured by direct observation time study differs substantially. This leads us to the conclusion that it is even necessary to analyze manual work procedures in VR by means of predetermined times, such as MTM-2, since the overall completion time measured by direct observation is higher in VR.

7 Outlook

Future work should supplement user studies by standardized questionnaires such as the NASA TLX [12] for measuring additional task loads that might be evoked by using a VR system. These studies could address a more complex work scenario at a given industrial workplace, employing professional workers to participate.

While the detection and transcription of basic motions was done manually for our study, the full potential of an MTM analysis in VR can be harnessed with an automated transcription. Four days were needed for the manual MTM-2 transcription of 21 users' recordings, both in VR and reality, with two trials each. This potential time saving through automated transcription can be considered as a rough estimate for the actual savings potential. Therefore, our study can also be seen as a proof of concept for a fully automated MTM analysis, which was proposed by Bellarbi et al. [2].

This leads us to the next research question about the number of trackers and their position on a user's body for capturing body motions. In our study, we track the head position together with one hand-held controller. However, it is likely that this will not be sufficient to identify all body motions and poses, e.g., bending. We agreed that this issue could be solved by attaching trackers to the user's pelvis or feet. Further studies could focus on the optimal number of trackers and their placement or the most suitable VR hardware setup to capture and further analyze body motions.

Another issue for the automated transcription could be the separation of intended body motions that are required for completing the task from unintended body motions for

handling the VR system itself. Unintended body motions could be caused by the uncomfortable attachment of the HMD or an interfering cable required by the VR system.

In order to move towards an even more detailed analysis such as MTM-1, we envision to also technically capture precise hand gesture detection. There are different ways to recognize gestures, but the most common ones used in VR are optical and inertial tracking. For optical tracking, systems like the Oculus Quest II could be used. Inertial tracking for hand gesture recognition usually consists of equipment with some mounted inertial trackers, e.g., Sensoryx gloves. It may be possible that using inertial gloves or optical tracking instead of controllers will be less accepted by the user since haptic feedback cannot be addressed yet, and thus grasping a virtual object may feel unnatural.

Acknowledgements We further want to thank Inspacion AG for providing the virtual factory surroundings.

Author contribution

- V. Gorobets: Acquisition and analysis of the study data, paper writing.
 V. Holzwarth: Analysis of the study data, paper writing.
 C. Hirt: Paper writing.
 N. Jufer: Industrial advisor.
 A. Kunz: Project management, acquisition and analysis of the study data, paper writing.

Funding Open Access funding provided by ETH Zurich. This work is fully funded by the Swiss Innovation Agency *Innosuisse* as part of the Eurostars project with the number “E!113504.”

Availability of data and material The data is completely available.

Declarations

Ethics approval According to the ethics committee, there was no ethical approval necessary for the here presented study. The recorded data was properly anonymized and participants cannot be identified using the here presented data.

Consent to participate All of the participants were recruited on their own will and were informed that they could stop the study session for any reason at any given time.

Consent for publication All authors give their consent for publication of the here presented work.

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted

use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Andrianakos G, Dimitropoulos N, Michalos G, Makris S (2019) An approach for monitoring the execution of human based assembly operations using machine learning. *Procedia CIRP* 86:198–203. <https://doi.org/10.1016/j.procir.2020.01.040>
- Bellarbi A, Jessel JP, Dalto LD (2019) Towards method time measurement identification using virtual reality and gesture recognition. In: 2019 IEEE International conference on artificial intelligence and virtual reality (AIVR). IEEE. <https://doi.org/10.1109/AIVR46125.2019.00040>
- Benter M, Kuhlang P (2019) Analysing body motions using motion capture data. In: *Advances in intelligent systems and computing*. Springer International Publishing, pp 128–140. https://doi.org/10.1007/978-3-030-20040-4_12
- Bubb H, Engstler F, Fritzsche F, Mergl C, Sabbah O, Schaefer P, Zacher I (2006) The development of ramsis in past and future as an example for the cooperation between industry and university. *International Journal of Human Factors Modelling and Simulation* 1(1):140–157. <https://doi.org/10.1504/IJHFMS.2006.011686>
- Buzjak D, Kunica Z (2018) Towards immersive designing of production processes using virtual reality techniques. *Interdisciplinary Description of Complex Systems* 16(1):110–123. <https://doi.org/10.7906/indec.16.1.8>. <http://indec.eu/index.php?s=x&y=2018&p=110-123>
- Caiza G, Ronquillo-Freire PV, Garcia CA, Garcia MV (2021) Scoping review of the work measurement for improving processes and simulation of standards. In: *International conference on information technology & systems*, pp 543–560. https://doi.org/10.1007/978-3-030-68285-9_51
- Chan D (2003) Simulation modelling in virtual manufacturing analysis for integrated product and process design. *J Assem Autom* 23(1):69–74
- Chedmail P, Chablat D, Le Roy C (2003) A distributed approach for access and visibility task with a manikin and a robot in a virtual reality environment. *IEEE Trans Ind Electron* 50(4):692–698. <https://doi.org/10.1109/TIE.2003.814760>
- Chryssolouris G, Mavrikios D, Fragos D, Karabatsou V (2000) A virtual reality-based experimentation environment for the verification of human-related factors in assembly processes. *Robot Comput Integr Manuf* 16(4):267–276. [https://doi.org/10.1016/S0736-5845\(00\)00013-2](https://doi.org/10.1016/S0736-5845(00)00013-2)
- Gaisbauer F, Agethen P, Otto M, Bär T, Sues J, Rukzio E (2018) Presenting a modular framework for a holistic simulation of manual assembly tasks. *Procedia CIRP* 72:768–773
- Goos J, Lietaert P, Cools R (2021) Computer assisted ergonomic assembly cell design. *Procedia CIRP* 97:87–91. <https://doi.org/10.1016/j.procir.2020.05.208>. <https://linkinghub.elsevier.com/retrieve/pii/S221282712031427X>
- Hart SG, Staveland LE (1988) Development of nasa-tlx (task load index): Results of empirical and theoretical research. *Adv Psychol* 52:139–183. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- Honglun H, Shouqian S, Yunhe P (2007) Research on virtual human in ergonomic simulation. *Computers & Industrial Engineering* 53(2):350–356. <https://doi.org/10.1016/j.cie.2007.06.027>
- Jayaram U, Jayaram S, Shaikh I, Kim Y, Palmer C (2006) Introducing quantitative analysis methods into virtual environments for real-time and continuous ergonomic evaluations. *Computers in Industry* 57(3):283–296. <https://doi.org/10.1016/j.compind.2005.12.005>

15. Jianping J, Keyi H (2005) Development of an immersive virtual maintenance simulation system [j]. *Journal of Computer Aided Design & Computer Graphics*, 5
16. Kummelsteiner G (1994) 3d-bewegungssimulation als integriertes hilfsmittel zur planung manueller montagesysteme. In: iwv Forschungsberichte, vol 66. iwv
17. Kunz A, Zank M, Fjeld M, Nescher T (2016) Real walking in virtual environments for factory planning and evaluation. *Procedia Cirp* 44:257–262. <https://doi.org/10.1016/j.procir.2016.02.086>
18. Liker JK (2006) *Der Toyota-Weg: 14 Managementprinzipien des weltweit erfolgreichsten Automobilkonzerns*. FinanzBuch Verlag
19. Liu JH, Hou WW, Zhang ZX, Liu WD, Jiang K (2011) Virtual assembly technology based on precision and physical attribute. *Computer Integrated Manufacturing Systems* 17(3):595–604
20. Ma L, Zhang W, Fu H, Guo Y, Chablat D, Bennis F, Sawanoi A, Fugiwara N (2010) A framework for interactive work design based on motion tracking, simulation, and analysis. *Human Factors and Ergonomics in Manufacturing & Service Industries* 20(4):339–352. <https://doi.org/10.1002/hfm.20178>
21. Mühlstedt J, Kaussler H, Spanner-Ulmer B (2008) The software incarnate: digital human models for cax and plm systems. *Zeitschrift fuer Arbeitswissenschaft* 62(2):79–86
22. Müller B., The DN, Dang QV, Minh Duc B, Seliger G, Krüger J., Kohl H (2016) Motion tracking applied in assembly for worker training in different locations. *Procedia CIRP* 48:460–465. <https://doi.org/10.1016/j.procir.2016.04.117>
23. Osterlund J, Lawrence B (2012) Virtual reality: Avatars in human spaceflight training. *Acta Astronautica* 71:139–150. <https://doi.org/10.1016/j.actaastro.2011.08.011>
24. Qiu S, He Q, Fan X, Wu D (2014) Virtual human hybrid control in virtual assembly and maintenance simulation. *Int J Prod Res* 52(3):867–887. <https://doi.org/10.1080/00207543.2013.842025>
25. Schmidt M, Wendt A (1989) Arbeitskreisbericht kleingerätemontage. In: *Interner Bericht des Instituts für Werkzeugmaschinen und Betriebswissenschaften*
26. Seth A, Vance JM, Oliver JH (2011) Virtual reality for assembly methods prototyping: a review. *Virtual Reality* 15(1):5–20. <https://doi.org/10.1007/s10055-009-0153-y>
27. Stich V, Blum M (2015) A cybernetic reference model for production systems using the viable system model. In: Umeda S, Nakano M, Mizuyama H, Hibino N, Kiritsis Dimitris and von Cieminski G (eds) *Advances in production management systems: innovative production management towards sustainable growth*, IFIP Advances in information and communication technology, vol 459. Springer International Publishing, pp 169–176. https://doi.org/10.1007/978-3-319-22756-6_21
28. Wu D, Yang R, Ma D, Fan X (2008) Integrated virtual assembly environment and its application in ship piping layout. *Int J Production Res* 46(17):4729–4749. <https://doi.org/10.1080/00207540601008424>
29. Yang J, Pitarch EP, Kim J, Abdel-Malek K (2006) Posture prediction and force/torque analysis for human hands. Tech. rep., SAE Technical Paper. <https://doi.org/10.4271/2006-01-2326>
30. Yang J, Sinokrot T, Abdel-Malek K (2008) A general analytic approach for santos™ upper extremity workspace. *Computers & Industrial Engineering* 54(2):242–258. <https://doi.org/10.1016/j.cie.2007.07.008>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.