The impact of augmented reality on learning curves and mental workload: a preliminary experimental study

Maurizio Faccio*
maurizio.faccio@unipd.it
Department of Management and Engineering, University of Padua
Vicenza, Italy

Irene Granata*
irene.granata@phd.unipd.it
Department of Management and Engineering, University of Padua
Vicenza, Italy

Leonardo Maretto*
leonardo.maretto@phd.unipd.it
Department of Management and Engineering, University of Padua
Vicenza, Italy

ABSTRACT

The learning process has always been fundamental in the industrial environment to correctly learn the right process and to perform it faster, increasing efficiency by also minimizing errors, consequently. Nowadays, new technologies that are emerging in this field are based on augmented reality, and, through a motion capture architecture, it is possible to real-time follow the operators' activities and guide them in the next ones, improving the learning process. Therefore, this paper presents an architecture setup and the first preliminary tests, realized in the Industrial Plants and Logistics Laboratory of the University of Padua, in order to study the benefits that this type of technology can provide. The main findings are the decrease in the time required to learn the job along with a smaller operator's cognitive workload during the training.

KEYWORDS

Motion capture, Tracking, Learning, Augmented reality, Information retrieval

ACM Reference Format:
Maurizio Faccio, Irene Granata, and Leonardo Maretto. 2023. The impact of augmented reality on learning curves and mental workload: a preliminary experimental study. In *2023 9th International Conference on e-Society, e-Learning and e-Technologies (ICSLT 2023), June 09–11, 2023, Portsmouth, United Kingdom.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3613944.3613950

1 INTRODUCTION

The learning curve (also called learning or progress) is a tool that can be used both for strategic assessments related to productive competitiveness and to design (or reorganize) production systems taking into account changes that occur over time as a result of the learning phenomenon. In particular, the learning curve indicates the relationship between the time required for learning and the amount of information correctly learned; this is because we gain experience each time we repeat some activity. The curve, in particular, is described as a general improvement in performing a task, thanks to the repetition of the exercise over time [12]. The first and most important study on the topic was conducted by [2], who proposed the model known as "Power Curve" which can be applied to a wide variety of applications. The basic form of the model is:

$$t_n = t_1 \cdot n^{-b} \tag{1}$$

where n is the number of cycles completed, t_n is the performance time to complete the n^{th} cycle, t_1 is the performance time to complete the first cycle and b is the learning constant. This means that it is possible to obtain the *Learning Slope* Φ as follow:

$$\Phi = 100 \cdot 2^{-b} \tag{2}$$

This value is defined as the percentage of performance reduction, each time the experience is doubled and it represents the *rate* of learning, and in typical industrial applications is about $65 - 92\%$. In particular, the higher the value the less demanding and time spending it is to learn the task Nowadays, since the migration from Industry 4.0 to Industry 5.0 with its principles of operators' well-being, new technologies are emerging [4], also for the training of the operators in industrial applications, i.e., the process of learning the tasks to be performed, and, for this reason, it can be helpful to perform an evaluation of the benefits that they can offer in the learning process. One of these emerging technologies can be developed through the use of Motion Capture (MoCap), which is the process of live capturing the body movements of a subject, in order to translate them into a mathematical model that can be used as input for other systems [9]. Analyzing the state of the art, it is possible to find different examples of camera-based markerless motion capture systems applied in the industrial sector [5, 8, 13]. Other works also used feedback, given with augmented reality systems, to the operators [1, 3, 10]. Among these contributions, there are few that presented an integration of Motion Capture systems in the learning process. A Smart-Assembly-Workplace was proposed by [7], in order to transfer the knowledge to operators with very low experience. Another solution was proposed by [6] where through the use of a neural network the authors realized an assistive setup for assembly. However, because of the lack of contributions on how these systems influence performance, this paper proposes the implementation of a complete experimental setup, that combines depth cameras and skeleton tracking software, to evaluate how the integration of these technologies, compared to the traditional supports. The first preliminary results, reported in this paper, show that the learning process is improved, reducing both the cycle time and the number of errors. Moreover, the analysis was also focused on the evaluation of a human factor: the mental workload. The paper is organized as follows: Section 2 is for the proposed experimental setup, Section 3 established which variables

*All authors contributed equally to this research.

ICSLT 2023, June 09–11, 2023, Portsmouth, United Kingdom

ACM ISBN 979-8-4007-0041-5/23/06.
https://doi.org/10.1145/3613944.3613950

are considered in the tests, while Section 4 presents the results. Finally, Section 5 draws the conclusions and the future agenda.

2 EXPERIMENTAL SETUP

The purpose of the following experimental setup is to compare the learning curves that can be obtained with two different supports, in a pick&place application. The application consists in moving some objects from one pallet to the other a certain number of times, with two different types of support systems: *traditional instruction-based*, where the task order to be followed by the testers is represented by an ordered table displayed on a screen, and *augmented reality-based*, where the task order is represented by a set of prompts and suggestions displayed on the same screen. In the latter case, the testers do not know the sequence in advance, and the pick and place positions are shown in real-time: the next task to be performed is shown on the monitor only if the previous task has been correctly completed. To implement this support, Motion Capture (MoCap) technologies are necessary because it is required to know in real-time the positions of the hands of who is performing the tasks since a pick&place activity is associated with two specific positions in the pallets. g. The motion capture architecture proposed includes an *Intel RealSense D435* camera, and it uses an RGB sensor and two sensors for stereophotogrammetry that can measure the distance of a point from the position of the camera. Motion capture is done through *OpenPose* library, which is used for body joints position recognition in real-time. As stated before, the experiment consists of pick&place activities, meaning that it is required to move some objects from one pallet to the other. In order to collect all the necessary data, i.e., pick time, place time, total time, and errors, two smart pallets are designed. These smart pallets are wooden boxes that include jigs to rapidly change the type of object because the tests are carried out with two types of objects, i.e., small spheres and small cubes, that have to be inserted in the corresponding holes. In the lower plate, moreover, beneath each hole, proximity sensors are installed. The use of these smart pallets guarantees not to be tied to the specific process chosen for the assessment of learning. Typically, in fact, the times and learning factors obtained are constrained by the specific operations that must be performed. The setup is also integrated with the 120Hz binocular Pupil Lab eye tracker, which allows monitoring the level of mental workload reached by the tester, through the measurement of the blinks [11]. In particular, the blinks rate, and latency is analyzed. The complete test station is shown in Figure 1.

3 EXPERIMENTAL VARIABLES

In order to carry out the tests, different variables are considered and summarized in Table 1. The difficulty of the task is based on the

Table 1: Experimental variables

Variables	Levels	
Difficulty of the task	Simple	Complex
Number of Pick&Place	4	9
Type of support	Traditional instruction-based	Augmented reality-based

level of required mental workload, and it is modified through the use of different objects: small spheres that do not require orientation

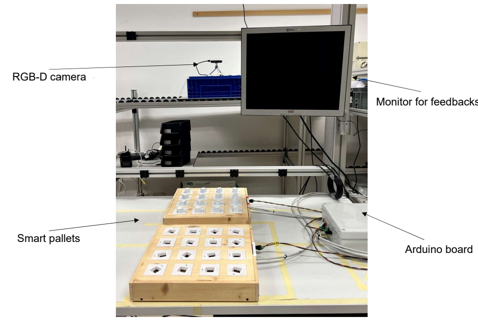
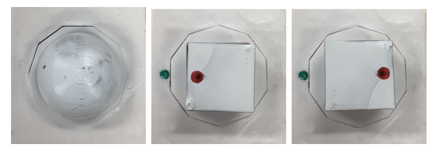


Figure 1: Experimental setup



(a) No orientation **(b) 0° orientation** **(c) 180° orientation**

Figure 2: Objects for the tests

(simple task), Figure 2a, and small cubes that require orientation (complex task). To specify their correct positioning, the cubes have a red circle that has to be placed in correspondence (0°) or at 180° with respect to the green signs on the jigs; this means that for the cubes there are two possible orientations. An example is shown in Figures 2b, 2c. In the test, half of the cubes have to be placed at 0° and the other half at 180°, following the specified instructions. The number of tasks to be performed is 4 or 9, meaning that for each test the participant has to pick and place four or nine objects with a specified sequence. The type of support is *traditional instruction-based* or *augmented reality-based*, as described above. Each participant has to perform only one level of difficulty but with all the combinations of the remaining variables, e.g., 4 spheres traditional instruction-based, 4 spheres augmented reality-based, 9 spheres traditional instruction-based, and 9 spheres augmented reality-based.

4 PILOT TESTS AND PRELIMINARY RESULTS

Thus far, 12 pilot tests were conducted, in particular, 6 participants tested the spheres task while the other 6 tested the cubes task. For the tests with 4 objects (both spheres and cubes) a group of 15 repetitions of the sequence was requested, while for the tests with 9 objects 20 repetitions of the sequence were requested. For each group of tests, the imposed sequence was different in order not to have a previous training level derived from the earlier tests. The number of tests is mathematically defined following the learning curves, the *Most* approach [14], for the definition of the standard times, and the learning factors from the literature. The preliminary results are shown in Table 2, divided by the type of support used, i.e., *traditional instruction-based (TI-B)* and *augmented reality-based (AR-B)* and by the type of object used. Moreover, as exempli gratia, in Figure 3 and in Figure 4 are shown the learning curves of two

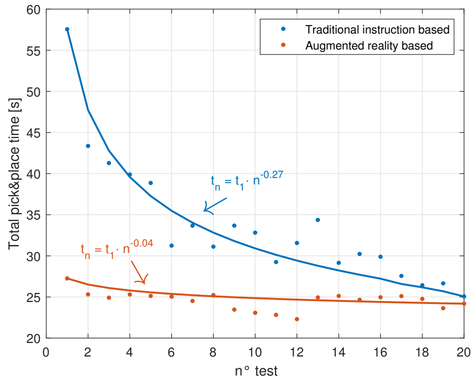


Figure 3: Tests with 9 spheres - 1^{st} participant

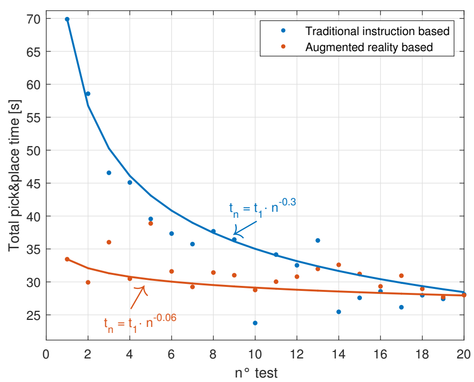


Figure 4: Tests with 9 cubes - 2^{nd} participant

tests. These figures are derived from the test with 4 objects and from the test with 9 objects, showing on the *x-axes* the repetition number and on the *y-axes* the total pick&place time for each repetition.

As it can be seen from the table, for the tests with both the supports, the curves obtained have a b factor between 0.35 and 0.04, meaning a Φ between 78% and 97%. It is also understandable that for the tests with the *AR-B* support, typically, the rates of learning Φ obtained are higher than the values obtained with the *TI-B* support, meaning that it is easier to learn the correct sequence and to perform it in a lower overall time. The quite interesting result is related to the fact that, if the tests with 9 objects are considered, the improvement in the learning, thanks to the *AR-B* support, is generally 10% higher than the learning with 4 objects, meaning that the more challenging the test, the more this type of support can be of benefit. Moreover, the total number of errors obtained for each test, with the second type of support is obviously equal to *zero*, since if the tester does not perform the correct activity the next suggestion is not generated, while for the other support the number of errors obtained is reported in Figure 5. Starting from this analysis, it is possible to evaluate the cases in which the experimental values of the time to complete the first cycle, t_1, coincide with the theoretical values. From the literature [14],

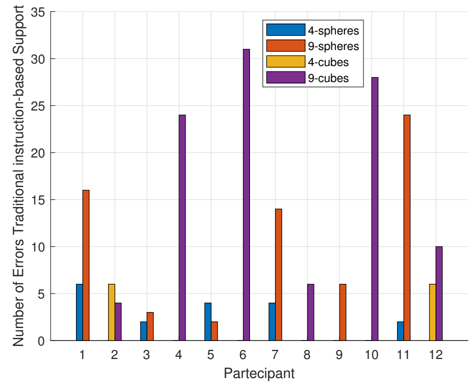


Figure 5: Total number of errors *traditional instruction-based* support

in order to define t_n, i.e, the performance time to complete the n^{th} cycle, the *Basic Most* method can be used. In particular, the pick&place activity is divided into the phases of *Get (G)*, *Put (P)*, and *Return (R)*, assigning a rating for each of them. The values obtained are G=1, P=1, and R=1, meaning a total of 30 TMUs (Time Measurement Units) since the sum has to be multiplied by 10 as the method imposes. Each TMU corresponds to 0.036 s, meaning that the standard time to move one object is 1.08 s. From this, the time for moving 4 objects is $C_4 = 4.32\ s$, while the time for moving 9 objects is $C_9 = 9.72\ s$. From [2], t_1 can be evaluated as in Eq. 3:

$$t_1 = (53.68 - 0.57\Phi) \cdot C \tag{3}$$

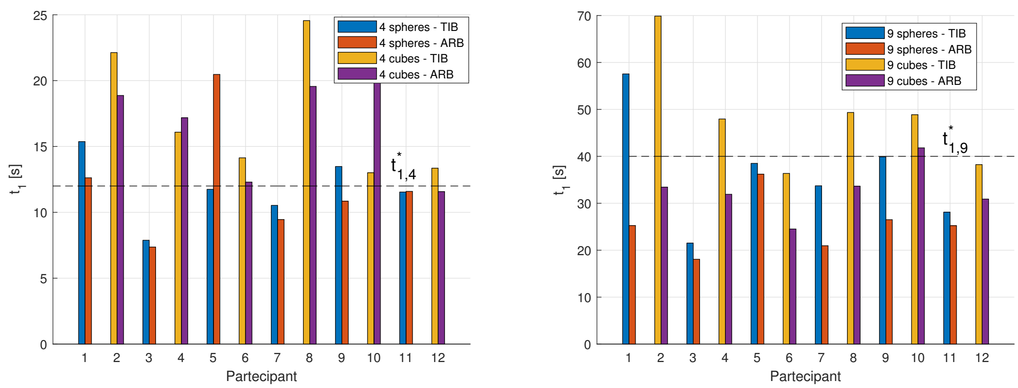
while the typical learning slope for activities like the tested ones, is $\Phi \simeq 90$ for the simple task and $\Phi \simeq 87$ for the complex one, meaning $t_{1,4} \simeq 12\ s$ and $t_{1,9} \simeq 40\ s$. The values obtained for t_1 are shown in Figure 6a and in Figure 6b respectively. In both cases, the times obtained with the *traditional instruction-based* support are worse than the ones obtained with the *augmented-based support*, meaning that, already from the beginning, the latter support can help to obtain better performance. From now, focusing only on the 9 objects tests, the differences in the learning rates can be analyzed, taking into account the last row of Table 2, where the average values of b and Φ are reported. In particular, the difference is defined as in Eq. 4, where C=cubes, S=spheres.

$$\Delta\Phi = \Delta\Phi_C - \Delta\Phi_S \tag{4}$$

where $\Delta(\cdot) = \bar{(\cdot)}_{AR} - \bar{(\cdot)}_{TI}$. In particular, the values obtained are:$\Delta\Phi_S = 6.89$ and $\Delta\Phi_C = 10.43$, meaning $\Delta\Phi = 3.53$, i.e., the increase in the learning slope is bigger for the cubes tests thanks to the *AR-B* support, implying that this type of support can provide more benefits in the more complex task during the learning process. Moreover, the preliminary results, for the 9 objects tests, of the eye-tracking analysis, tested by the last six subjects, are here reported: Figure 7a is for the blinks rate during the tests and, Figure 7b is for their latency. The meaning of these parameters is reported in Table 3, where the increase in the mental workload is associated with the symbol in the third column; e.g., if the blinks rate increases the mental workload decreases. From the figures, it is understandable

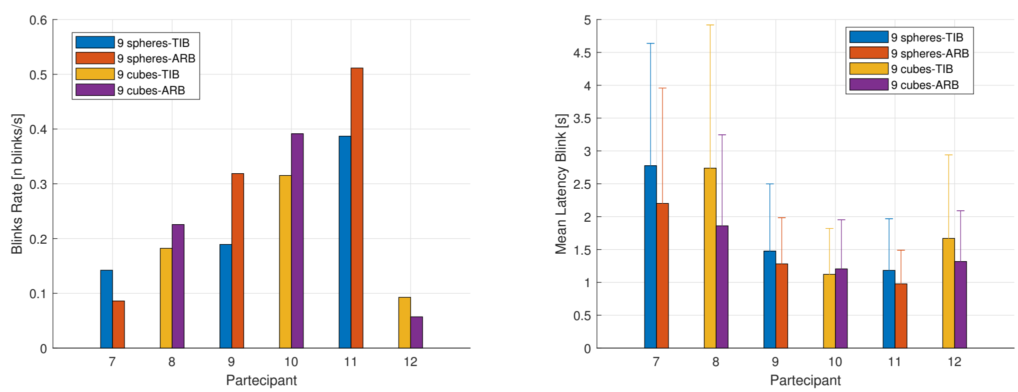
Table 2: Learning factors of the pilot tests

Participant	4 spheres				9 spheres				4 cubes				9 cubes			
	TI-B		AR-B		TI-B		AR-B		TI-B		AR-B		TI-B		AR-B	
	b	Φ [%]	*b*	Φ [%]	*b*	Φ [%]	*b*	Φ [%]	*b*	Φ [%]	*b*	Φ [%]	*b*	Φ [%]	*b*	Φ [%]
1^{st}	0.20	87.06	0.11	92.66	0.27	82.93	0.04	97.27	-	-	-	-	-	-	-	-
2^{nd}	-	-	-	-	-	-	-	-	0.35	78.46	0.11	92.66	0.30	81.23	0.06	95.93
3^{rd}	0.13	91.38	0.06	95.93	0.10	93.30	0.04	97.27	-	-	-	-	-	-	-	-
4^{th}	-	-	-	-	-	-	-	-	0.35	78.46	0.25	84.09	0.25	84.09	0.10	93.30
5^{th}	0.25	84.09	0.22	85.86	0.20	87.06	0.10	93.30	-	-	-	-	-	-	-	-
6^{th}	-	-	-	-	-	-	-	-	0.10	93.30	0.15	90.13	0.27	82.93	0.04	97.27
7^{th}	0.22	85.86	0.10	93.30	0.15	90.13	0.20	87.06	-	-	-	-	-	-	-	-
8^{th}	-	-	-	-	-	-	-	-	0.30	81.23	0.20	87.06	0.25	84.09	0.06	95.93
9^{th}	0.30	81.23	0.15	90.13	0.25	84.09	0.03	97.94	-	-	-	-	-	-	-	-
10^{th}	-	-	-	-	-	-	-	-	0.20	87.06	0.15	90.13	0.20	87.06	0.10	93.30
11^{th}	0.25	84.09	0.20	87.06	0.25	84.09	0.15	90.13	-	-	-	-	-	-	-	-
12^{th}	-	-	-	-	-	-	-	-	0.20	87.06	0.15	90.13	0.20	87.06	0.10	93.30
Average	0.23	85.62	0.14	90.82	0.20	86.93	0.09	93.83	0.25	84.26	0.17	89.03	0.25	84.41	0.08	94.84



(a) t_1 values for 4 objects tests **(b) t_1 values for 9 objects tests**

Figure 6: t_1 values obtained compared with the standard ones



(a) Blinks rate **(b) Blinks latency**

Figure 7: Blinks evaluation as mental workload driver

that the blinks rate is overall smaller with the *TI-B* support, meaning a higher mental. The same result is confirmed by the bigger latency. Studying the $\Delta(s)$, as done for the learning factors Φ, some considerations can be made. In particular, the differences in the average values of the blinks rate (R) and the latency (L) between

Table 3: Relation of blinks measures and mental workload

Measure		Mental Workload
Blinks	Rate	-
	Latency	+

the two types of supports are analyzed, as described by Eq. 5, 6.

$$\Delta R = \Delta R_C - \Delta R_S \tag{5}$$

$$\Delta L = \Delta L_C - \Delta L_S \tag{6}$$

where $\Delta(\cdot)_{C,S} = \bar{(\cdot)}_{AR} - \bar{(\cdot)}_{TI}$. For the blinks rate, the values obtained are $\Delta R_S = 0.066$ $blinks/s$ and $\Delta R_C = 0.028$ $blinks/s$, that means $\Delta R = -0.0379$ $blinks/s$, while for the latency the values obtained are: $|\Delta L_S| = 0.325$ s and $|\Delta L_C| = 0.382$ s, that means $\Delta L = 0.058$ s. These last are considered in absolute value because of the definition of the $\Delta(s)$; in fact, since the blinks latency of the *AR-B* support is smaller (driver of less mental workload) the difference with the blinks latency of the *TI-B* support results negative. From these results, it appears that in both cases the *AR-B* support is helpful in the reduction of the mental workload during the tasks, but from the blinks rate analysis, it seems to offer bigger support when the tasks are easier while from the blinks latency analysis, it seems to be more useful when the tasks are more complex. Table 4 summarizes all the $\Delta(s)$ evaluations. Generally, it can be concluded

Table 4: $\Delta(s)$ evaluations

	$\bar{\Phi}$	$\bar{R}$	$\bar{L}$
$\Delta(\cdot)_S$	6.89	0.066	-0.325
$\Delta(\cdot)_C$	10.43	0.028	-0.382
$\Delta\Phi$	3.53		
ΔR		-0.0379	
ΔL			-0.058

that with the *augmented reality-based* support the mental workload, both for the spheres and the cubes, with 4 and 9 objects as well, is lower, meaning less effort is required during the tests.

5 CONCLUSIONS

Nowadays, the arising of new technologies has also led to the development of new supports for the learning process because of its importance in the industrial environment, in order to integrate the wellness principles of Industry 5.0. Indeed, the better the support in the learning phase, the faster this is, and with a lower tendency to perform errors, resulting in an increment in both productivity and efficiency. For this purpose, i.e., to improve the learning phase, an architecture setup based on Motion Capture systems is here presented, with the aim to realize an *AR-B* support that guides the operators during the training. The main characteristic of this setup is the possibility of real-time following the operators' movements, and instructing them through video feedback in performing the correct activities, since it is always known which one they are executing and if they are making mistakes. Moreover, the use of smart pallets gives the opportunity to obtain results not linked to the timing of a specific process. In addition, the use of eye-tracker device consents to tracking the level of mental workload reached by the testers. From the preliminary results of the pilot tests conducted, the learning rates Φ, with this new support, are higher than the ones obtained with the traditional one, i.e., *TI-B* support, meaning less effort in the training phase. Moreover, as the complexity of the task increases this detachment is increasingly evident, both in terms of learning factors and in the time spent to complete the whole job, showing that this new kind of assistance can bring significant benefits. This is confirmed also by the analysis of the initial times t_1 that are significantly smaller with the *augmented reality-based* support than with the *traditional instruction-based* support. Moreover, the blinks rate is bigger, while their latency is smaller, resulting in a decrease in the mental workload. These are just the initial results, needed to provide an idea of the actual usefulness of these supports. The next agenda, of course, will be the continuation of the experimental tests, defining an adequate group divided by gender, age, level of education, and experience with this new support, in order to evaluate a representative sample of what is the industrial reality. Final results should be published by the end of 2023.

REFERENCES

[1] Michela Dalle Mura, Gino Dini, and Franco Failli. 2016. An integrated environment based on augmented reality and sensing device for manual assembly workstations. *Procedia Cirp* 41 (2016), 340–345.

[2] Ezey M Dar-El. 2013. *Human learning: From learning curves to learning organizations*. Vol. 29. Springer Science & Business Media.

[3] Maurizio Faccio, Emilio Ferrari, Francesco G Galizia, Mauro Gamberi, and Francesco Pilati. 2019. Real-time assistance to manual assembly through depth camera and visual feedback. *Procedia CIRP* 81 (2019), 1254–1259.

[4] Maurizio Faccio, Mauro Gamberi, Marco Bortolini, and Mojtaba Nedaei. 2018. State-of-art review of the optimization methods to design the configuration of hybrid renewable energy systems (HRESs). *Frontiers in Energy* 12, 4 (2018), 591–622.

[5] Florian Geiselhart, Michael Otto, and Enrico Rukzio. 2016. On the use of multi-depth-camera based motion tracking systems in production planning environments. *Procedia Cirp* 41 (2016), 759–764.

[6] Linh Kästner, Leon Eversberg, Marina Mursa, and Jens Lambrecht. 2021. Integrative object and pose to task detection for an augmented-reality-based human assistance system using neural networks. In *2020 IEEE Eighth International Conference on Communications and Electronics (ICCE)*. IEEE, 332–337.

[7] Bastian C Müller, The Duy Nguyen, Quang-Vinh Dang, Bui Minh Duc, Günther Seliger, Jörg Krüger, and Holger Kohl. 2016. Motion tracking applied in assembly for worker training in different locations. *Procedia CIRP* 48 (2016), 460–465.

[8] Michael Otto, Michael Prieur, Philipp Agethen, and Enrico Rukzio. 2016. Dual reality for production verification workshops: a comprehensive set of virtual methods. *Procedia CIRP* 44 (2016), 38–43.

[9] Dario Pavllo, Christoph Feichtenhofer, Michael Auli, and David Grangier. 2020. Modeling human motion with quaternion-based neural networks. *International Journal of Computer Vision* 128, 4 (2020), 855–872.

[10] Oliver Sand, Sebastian Büttner, Volker Paelke, and Carsten Röcker. 2016. smart. assembly–projection-based augmented reality for supporting assembly workers. In *International Conference on Virtual, Augmented and Mixed Reality*. Springer, 643–652.

[11] Da Tao, Haibo Tan, Hailiang Wang, Xu Zhang, Xingda Qu, and Tingru Zhang. 2019. A systematic review of physiological measures of mental workload. *International journal of environmental research and public health* 16, 15 (2019), 2716.

[12] Tom Viering and Marco Loog. 2022. The shape of learning curves: a review. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2022).

[13] YQ Wang. 2022. A Brief Analysis of the Development of Chinese Virtual Idol Industry Empowered by 5G+ Motion Capture Technology——Taking the Virtual Idol Group A-SOUL as an Example. In *Journal of Physics: Conference Series*, Vol. 2278. IOP Publishing, 012011.

[14] Kjell B Zandin. 2002. *MOST work measurement systems*. CRC press.