A Smart Factory in a Smart City: Virtual and Augmented Reality in a Smart Assembly Line

PETR HOŘEJŠÍ [ID], KONSTANTIN NOVIKOV, AND MICHAL ŠIMON

Department of Industrial Engineering and Management, University of West Bohemia, 30100 Pilsen, Czech Republic

Corresponding author: Petr Hořejší (tucnak@kpv.zcu.cz)

ABSTRACT Increasingly, the smart city space is requiring a reconceptualization of forms and factors of production, including factories and their place in the smart city space. Factories have always been a part of the city and many people spend a significant part of their lives there. Cities and factories share the same physical space and draw from the same resources, such as the energy grid, communication networks, public utilities, social connections, etc. Factories and cities should share the same IoT network in order to maximize their synergy level. In this view, as ICT-enhanced solutions are being implemented and so the concept of the smart city becomes a reality, it is mandatory that the connection between the smart city and the smart factory is examined. This paper represents the first step in this direction. We are presenting a new smart way to lighten the workload for employees (especially those involved in assembly, setup and maintenance) and increase factory efficiency. We have developed a brand-new smart solution for designing and presenting work instructions. The solution can be easily adapted to use in other fields like healthcare or smart-homes. This paper presents a comparison of different types of virtual/augmented and conventional assembly instructions. Today, we face the challenge of a lack of skilled employees and a high rate of employee turnover. Both result in huge time and production losses, because new employees have to be taught simple assembly tasks over and over again. In addition, as companies begin hiring many more foreign workers who do not understand the local language, the challenge of teaching becomes even more acute. Despite this, in modern production systems we can still find ineffective and complicated books and manuals with assembly, service and measurement instructions. We have prepared several variants for non-trivial multi-step assembly instructions: traditional ''paper'' instructions, video instructions, virtual instructions on screen (with/without in-situ projection and with/without a special controller). We have developed our own software system for working with and developing virtual assembly instructions. In this case the in-situ augmentation is a projection on to different parts of the workplace. 60 subjects were tested over two years in order to gather the learning curve for each of 5 types of instructions using virtual and augmented reality. We have proven that using any type other than ''paper'' will shorten the learning time by approximately half. Practitioner summary: We have prepared and tested variants for non-trivial multi-step assembly instructions. 60 subjects were tested over two years in order to gather the learning curve for all 5 types of instructions – traditional paper, video instructions, virtual instructions and two types of virtual instructions combined with augmented reality in-situ projection. We have proven that using any type other than ''paper'' will shorten the learning time by approximately half.

INDEX TERMS Assembly instructions, augmented reality (AR), virtual reality (VR), industrial ergonomics, training, virtual environments.

I. INTRODUCTION

The technological breakthrough enabled by the vision of the fourth industrial revolution is opening new

horizons of robotization and automation in production and beyond the factory [27]. The current modernization is being taken out of cities and the implementation of the digital economy is resulting in smart plants, companies and organizations which are using completely computer aided manufacturing [28]. On the other hand, some methods which had

The associate editor coordinating the review of this manuscript and approving it for publication was Miltiadis Lytras [ID].

been used previously only in the industrial sector are being applied in smart cities, such as study [29], where discrete event simulation was used to create designs for smart hotels.

Smart cities require a new, progressive way to carry out education and training [31]. We will focus on training and education using virtual animated step-by-step instructions. Although there have been advances in technologies that can be used to present instructions in general, traditional paper books with assembly, service or measurement instructions are still commonly found in modern production systems [1]. These are prone to getting damaged or lost, and they can be misunderstood by foreign employees.

In the European Union (and also countries outside it) there is a lack of employees and a high rate of employee turnover [24], causing huge losses in the form of time and erroneous production by always teaching new employees simple assembly tasks over and over again. The situation is further complicated by the fact that in most companies the training process is not standardized. This is also associated with the slow launch of new products, as traditional manuals are often cumbersome and take a long time to understand. It is usually foreign employees who do not understand the local language, requiring another worker to teach them individually.

We have been seeking a new, modern way to reduce the negative impact of these factors. Also, there is a rising tendency of employing cognitively impaired workers, who have special needs while they are working.

The idea of using digitalized assembly instructions is not new, but it is one of the main tools of the Industry 4.0 concept [22]. The area of Computer Aided Instructions (CAI) has been here for a while, but it can also mean just displaying the same information on a screen instead of printing it on paper. Such an approach provides improvements in maintaining the instructions, but the impact on work performance is not as high [2]. On the other hand, changing the text instructions for static CAD images with motion diagrams improves the learning rate, but animating the CAD models improves it even further [3].

Electronic devices enable these instructions to be 3D animated using virtual reality (VR) or augmented reality (AR). From the research of various implementations and experiments, these can be divided according to the presentation mode into:

Instructions on screen: A computer monitor, tablet or data projector is present so that the instructions are displayed at a location near to the assembly task.

Instructions in a Head-Mounted Display (HMD): The instructions are displayed in the wearer's field of view. The main advantage is having free hands and the relative simplicity of the system. Care must be taken so that the instructions do not cover or interfere with the actual task by placing them to the edges or corners of the display or by tracking the user's view and displaying them beside the workplace. Such approaches are discussed in [2], [4], [5]. However, this type of instruction projection is not universally beneficial, and is useful mostly in confined situations where it is hard to reach a tablet or paper or when it is not possible to look away from the situation or when having both hands free is necessary. Cost, battery life, image quality, lighting conditions and fragility are also common factors inhibiting the deployment of HMDs. These displays also cause higher eye and mental strain [6].

Augmented Reality (AR) instructions: The instructions are registered with the actual task. Markers or computer vision algorithms are used to detect the object's position and to match the animated 3D models. Such systems are usually found as mobile phone applications or stationary cameras with a computer screen [21]. Deployment is also problematic in some cases because mobile devices will not work long enough due to low battery life, low processing power and poor cooling designs – computer vision algorithms are very demanding on computing power (this can be compensated by the use of a desktop computer for calculations, but at the cost of losing mobility, or solved hybridlike using a cloud system). Also, the benefits in time savings have not been proved. Some studies even found worse times compared to paper, although the error rate is usually lower [7], [8].

AR instructions in an HMD: This solution is an ideal approach for presenting any kind of instructions. Currently, there are no mass deployment ready solutions. Basically, the technical challenges mentioned in the previous two paragraphs apply here, adding the requirement of calibration of the user's vision with the screen. Efficiency of an AR HMD has been claimed in [6] and [8] in picking tasks (higher speed and lower error rate). [2], [7] and [9] found different, almost opposing, results for assembly tasks. This applies for optical see-through glasses, whereas video see-through glasses have been proven to be inferior [10], [11]. However, welding tasks can profit from video see-through HMDs [12].

In-situ projected instructions: In-situ projection means projecting the information in the place of the actual task. For example, a projector can be installed above the task and project the instructions directly. Funk's research [2] has proven the benefits of such projections, but they also prove that a negative effect can occur when workers already know how to conduct the assembly task. It is probably a distraction, as skilled workers were found to have about 10% worse performance [13], [14]. [15] describes using laser projectors for location highlighting when, for example, drilling holes.

Augmented reality still has major drawbacks today. The limitations are transparent projection quality, calibration of the user's view and object recognition and tracking. Finally, wearable devices require too much computational power. In 2004 and 2011 there were research studies [16], [17] about the lack of so called "killer applications" in the field and it is true to say that the situation still remains. Also, to evaluate the real benefits, we must wait for a truly high-quality AR device to make experiments on.

Because of this, we have been focusing on ways to replace conventional paper manuals with 3D animated instructions on a screen placed in the workplace. Therefore, a software tool

to make these easily and without any special skills is needed, as well as experience and best practices.

The aim of this paper is to demonstrate the difference of times when learning an unknown assembly task from paper and from other kinds of virtual assembly instructions. The outcome will either prove or disprove the assumption that 3D animated assembly instructions have a significant impact on shortening the time of the training process.

This work stems from a previous work [26], where a pilot software solution was also tested on a group of probands.

II. DEVELOPMENT

Although it is not exactly a challenging task to make animated instructions, things get technical when making a toolkit for industrial or mechanical engineers to let them make the instructions quickly and easily.

In order to have a variable custom solution, in which we will be able to test various technical options, we have created our own software tool for creating and presenting work instructions in VR and AR. This software solution is prepared for VR realized using a Head-Mounted-Display or a CAVE (Computer Automated Virtual Environment) and also for AR using marker or markerless tracking presented on a monitor, cell-phone or see-through device or AR using In-Situ projection. Because of the current state hardware limitations (discussed in the conclusion), we have chosen for this study only technical options, which are directly and easy implementable in a current state-of-the-art Smart Factory – VR presented on a display and AR presented using In-Situ projection (with several variants).

The relatively common data structure, consisting of animation steps and movements within these steps has been described for example in [18]. Our data structure adds a parenting root. The whole animation should then be played in the local coordinate system of this root.

This data structure must be further enhanced for AR compatibility. The root is attached to a marker (e.g. an image marker or recognized object, etc.). A challenging task is that if the object changes through the assembly process, the object recognition markers will stop working within several steps. On the other hand, when mounting small parts into a big frame, the marker can stay constant. But we are seeking a universal solution, and that is why we add marker data to each step. The markers will be aligned, and a root attached to them.

Each animation step has a set of objects that are static within the step: a frame, already assembled parts, etc. Some of these do not have to be displayed when AR tracking is present – they are already there, there is no need to display them. We divide these static objects into two groups: the first to be displayed only in non-AR mode, and the second to be displayed all the time.

Attention should be paid to the actual types of animations. We categorize the animations to "outer" and "inner". The "outer" category are animations that can be animated using a transformation matrix on the whole 3D object – translation, rotation and scaling. The "inner" category means animations
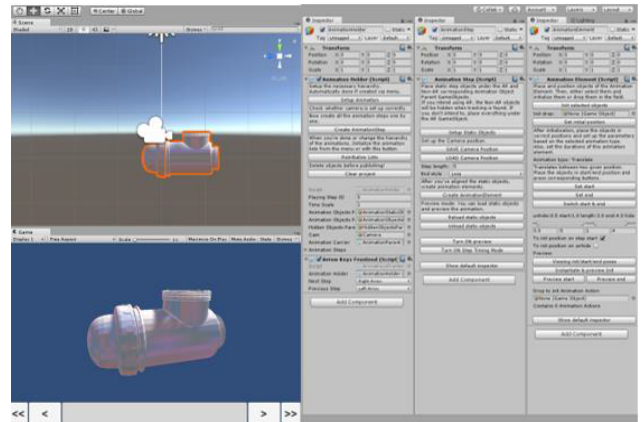
transforming the object in a more complex way, driven by a function of interpolation between two states – f(t) where t is between 0 and 1. Examples are flashing colors, segmented bending or interpolation between two animation key frames.

For implementation of the first prototype, we chose Unity3D because it has a powerful graphics core and capabilities to utilize various VR and AR devices and toolkits. It is also capable of creating modal windows and applications within the Unity Editor, allowing for easy use. However, even Unity has some drawbacks. The main one (for us) is no uncommercial(free) support for CAD models that need to be converted into 3D mesh models.

A prototype of a VR/AR compatible assembly instructions editor has been developed and tested (see Figure 1). We have tested it in experimental conditions in order to evaluate the improvement of the instructions on screen versus paper instructions. We unfortunately could not test real AR assembly instructions as there is no technology available to us which is capable of markerless tracking of these very symmetrical and simple objects with a lack of contrasting geometric features.

Our system has also been tested in real production conditions. The methodology of the experimental testing is described later in this article, followed by a section on the experience acquired in a real-life scenario.

III. EXPERIMENT DESIGN

The goal of the experiment was to find out how the times differ between learning an unknown assembly task from paper and from other kinds of assembly instructions. A sink water trap was chosen as the reference object of the assembly task because of its good availability in small numbers, the high probability that it is unknown to the recruited test subjects and that the assembly instructions can be animated using two elementary kinds of motion: translation and rotation. We can consider this part as a complex part in accordance with a study [25]. For the assembly, a realistic test environment in the form of an assembly desk was created with a 19" LCD screen for the instructions – see Figure 2. The parts were in boxes

FIGURE 2. Workplace designed for the experiment, in configuration for group 3DIO (3D animated manual, in-situ marked box, special controller).

FIGURE 3. Paper instruction assembly preview (original in Czech language).

marked with part numbers, which consisted of S followed by a random 3-digit number. The workplace was ergonomically designed and optimized. The projector was mounted above the workplace.

First, a paper version of the assembly instructions was created (see Figure 3). Its accuracy and comprehensibility were tested on five volunteers. After testing by unbiased volunteers, a few small changes were made, and the entire manual was completed for subsequent testing. Before the water trap assembly could begin, each tested person had the task of locating a specific assembly instruction in a folder, which contains other 27 different assembly instructions. Each cover page looked the same.

The paper manual consists of a total of 16 pages. The home page explains that it is a workflow for the assembly of the water trap type A441P supplemented with a picture of the final product. On the second page is a photograph of

the assembly workplace and on the other pages there is a list of necessary parts, tools and a description of how to operate an electric screwdriver. Furthermore, the assembly instructions describe the steps that are accompanied by pictures. The assembly process has a total of 20 steps.

The other examined option is a video tutorial. Creating the video tutorial can be divided into three steps. The first step was to film all the assembly operations necessary to complete the water trap. In the second step, these videos were edited and divided into a total of ten complete assembly steps. In the last step the final manual was compiled in the form of a PowerPoint presentation. The video was recorded directly at the assembly site. The camera was placed on a tripod, on the edge of the mounting plate. This shot appeared to be the best shot and eliminated image stabilization problems. The scene is located relatively close to the camera and provides a detailed view of the assembled parts.

After completing and filming all the assembly scenes, all videos had to be cut and divided into individual steps for the final tutorial. Overall, the assembly takes place in ten steps. In the presentation there are ten frames with a video and the last frame is only a photograph of the final product. This image is used to check if the water trap has been assembled correctly and if they have forgotten anything. The animated installation instructions for the water trap were divided into eleven steps. In the Unity 3D Virtual Installation Guide, all eleven steps were created in the AnimationHolder folder. The steps were named 'Step' and numbered for quick reference. Each step is assigned part models that do not move in the animation and are only embedded in the scene.

First it is necessary to screw part S354 to S353, then it is necessary to insert part S381 with the narrow side into S353. Next, the S517 must be placed on the S381 part. S353 is screwed onto the prepared parts. In the next step the upper part of the water trap is assembled. The S085 part is inserted into the S305 part, then S090 is placed on it and finally the S116 part is added. Insert the screw S014 into the prepared system and screw it manually for 3-4 revolutions, using the electric screwdriver for the rest. This is triggered by pressing the switch to the lower position and after pressing on the desired part, in the case of insufficient pushing the screwdriver will not work. The screwed part must be held firmly. In the next step the seal S120 is inserted with the wider side into the part S048. Subsequently, the P050 part is inserted into the groove with the seal side, i.e. the black side down into the S363 part. Then the S048 component is screwed into the S363 part. In the next step, the seal S323 is pushed into the S363. The upper part of the S305 water trap is screwed on S363. S004, then S516 and finally S009 are pushed onto part S363. In the penultimate step, the other two assemblies are screwed together.

The steps still required us to select the appropriate position for the camera to capture the entire animation and the measure the time length of the step. After these necessary operations, animations were created for each step. The animations were again appropriately named, such as Shift S048, Screwing

FIGURE 4. Video assembly instructions preview.

FIGURE 5. Parts required for assembly.

S048, Inserting S120, so that in the event of an error, the user knows where to reach. Parts models that were not shown on the stage at that step are placed in the inactive GameObject folder and its Parts and Complete Parts subfolders.

To install the water trap, there are eighteen parts, and an electric screwdriver is required to tighten the S014 screw. An overview of the necessary parts for the step is shown in Figure 5. A button system was selected to control and switch between the steps in the assembly instructions.

IV. EXPERIMENT
The assembly instructions were tested on a group of university students aged 20 to 25 years (so we consider a consistent group). Participants filled in information on a form about their age, gender, education, height and manual skills, then they were instructed in how to operate the manual: how to start, switch and return to a given step, and how to shut down the manual after assembly. The group were also informed of the predefined conditions under which they will assemble the water trap. All participants were standing during the experiment, under artificial lighting from a fluorescent lamp above the workplace and at a room temperature of 21 °C.

The time between opening the first step (on screen, on paper, in the presentation) and completing the product was measured for each experiment (time for locating the right assembly instructions is not included).

The groups were tested using:

- PAP – paper instructions
- VID – filmed instructions
- 3D - animated instructions displayed on the screen
- 3DI - animated instructions displayed on the screen with in-situ marking of the correct parts box
- 3DIO - animated instructions displayed on the screen with in situ marking and controller with next and previous step buttons

The PAP group used only non-informatic standard tools. The VID group used a sophisticated video played on a conventional monitor. The other groups (3D, 3DI and 3DIO) used a 3D generated interactive scene projected on a monitor with a stereoscopic projection option. Additionally, 3DI and 3DIO groups used an augmented in-situ projection from a full HD data projector (resolution 1920 x 1080) hanging above them pointing down on the working table and racks. The projection was extended to the projector screen and generated using our own software. A similar projection is described e.g. in [37].

A total of 60 test subjects were recruited from the ranks of second grade mechanical engineering university students. There were 12 probands tested for each of five groups (i.e. $12 \times 5 = 60$). Because we need to estimate the learning curve, there is a new proband for each test. Experiments were conducted in spring 2017 and 2018. In 2017, two groups were tested. One group had paper instructions and the other had 3D animated instructions displayed on the screen (groups PAP and 3D). In order to fulfil the rigorous methodology, we needed to validate if 12 probands in a group is statistically enough just after the first tests. We compared PAP to VID data for the first iteration. Using ANOVA which was recounted to the Kruskall-Wallis test, the test strength is 94%. So, 12 probands for each group is statistically enough (if we would like to compare PAP with other variants).

In 2018, the assembly instructions were enhanced with in-situ marking of the correct parts box (group 3DI), another was supplemented by a specialized controller with next and previous step buttons (3DIO; this setup can be seen in Figure 2), and finally a group where the instructions were not animated but filmed (VID). The VID group had the correct box marked in the corner of the video. All group subjects assembled the water trap in 6 consecutive iterations (6 try-outs).

Data was collected directly using our own developed software using a software bridge to MS Excel in the case of 3D, 3DI and 3DIO. For PAP and VID a conventional calibrated timer was used. The following statistical analysis was performed in MathWorks Matlab.

V. EXPERIMENT RESULTS AND DISCUSSION
On the basis of the experiments a learning curve was created for each experiment - see Table 1 and Figure 6. The results

TABLE 1. Experiment results.

	1	2	3	4	5	6	Std. deviation	Average	Median
PAP	11:18	05:39	04:43	03:44	03:32	03:24	02:45	05:23	04:14
VID	05:44	03:52	03:13	02:53	02:44	02:38	01:04	03:31	03:03
3D	06:44	04:31	03:39	02:57	02:48	02:42	01:25	03:53	03:18
3DI	05:16	03:45	03:05	02:51	02:31	02:27	00:58	03:19	02:58
3DIO	05:38	03:42	03:23	02:46	02:24	02:19	01:08	03:22	03:04

TABLE 2. Correlation between methods.

	PAP	*VID*	*3D*	*3DI*	*3DIO*
PAP	1				
VID	0.993736	1			
3D	0.985864	0.997066	1		
3DI	0.982964	0.996452	0.996056	1	
3DIO	0.982542	0.989748	0.992583	0.992112	1
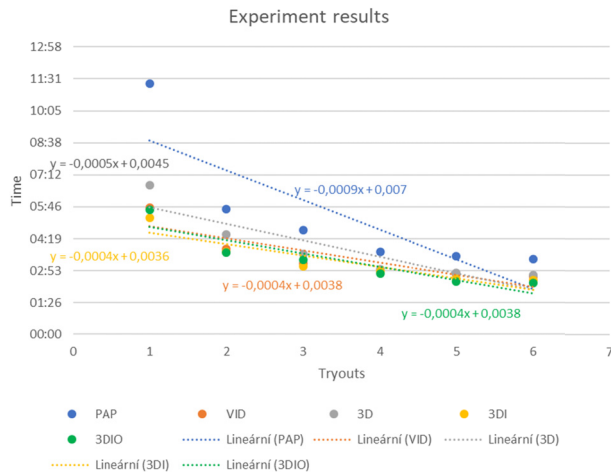
FIGURE 6. Experiment results.

show longer times in the PAP group. The PAP group was chosen as the reference. The learning curve shows how fast the workers work from the beginning when they are not yet familiar with the assembly process. Some results were discarded because some students made major errors requiring them to return a few steps back. Results where the students made minor mistakes which were corrected immediately were included in the final measurements, because we believe they could be encountered in the real environment as well. The average time taken to find the correct assembly manual in the bundle for PAP was 1 min 57 s (this time was not included in the following data).

From this we can see that the same instructions in the animated form yield 30% less time spent on the assembly task (group 3D) in comparison with the PAP group.

The other 3 groups yield very similar results of a nearly 40% time drop but are not very different from each other. In the 3D group, the correct parts were not highlighted, and the subjects had to find them by searching for the correct label. In the 3DI and 3DIO groups the boxes were highlighted directly; in the VID group the correct box was highlighted in a photograph which was displayed in the corner of the video. Overall, the VID group performed slightly worse in most iterations than the 3DI and 3DIO groups. Instead of clicking a button on the screen with the mouse, the 3DIO group subjects only pressed a button labelled Next (or Previous) on a device attached to the table. The differences are minor. We believe that with the average population, where computer literacy is not nearly as high as among university students,

the results would be more different. In manufacturing conditions, pressing buttons on a screen with a mouse-controlled cursor is considered a waste of time. It is interesting that with these test subjects, the 3DIO began with longer times and ended with shorter times compared to the 3DI group, which may implicate that the group is not familiar with uncommon computer equipment.

The results correlate with each other. The learning curves are therefore similar. The lowest correlation is in the first step between PAP and 3DIO.

Complete ''raw'' testing data served as the input for statistical data processing. The Jarque – Bera test with zero hypothesis was used to test the data: the data follows the normal distribution. Measurement of PAP, 3DIO and 3D instructions is based on the test according to the normal distribution (the null hypothesis was not rejected). Since the normal distribution in all cases has not been confirmed and there is not enough data to overlook this fact, the testing of the differences between the assembly instruction groups will be tested using the non-parametric Kruskall-Wallis test.

Therefore, we perform a non-parametric test to determine if there is a difference between the times of the five different types of instructions. We test the null hypothesis H_0,: Test groups (PAP,VID,3D,3DI,3DIO) have the same distribution function (no difference from each other) against the alternative hypothesis H_A: At least one group differs in its distribution function (at least one differs from the others).

If the null hypothesis is rejected (at least one group differs from the others), the so-called multiple comparison is carried out using the Tukey Kramer test, which can be used to determine which groups specifically have a statistically significant difference. So we have 5 groups and we test the null hypothesis about the agreement of the averages of two selected categories i,j H_0: $\mu_i = \mu_j$ as opposed to the alternative hypothesis of the average mismatch H_A: $\mu_i \neq \mu_j$, where i,j = 1,…n a i≠j. We always test pairs of groups with each other. We test all hypotheses at the significance level $\alpha = 5\%$.

Figure 7 shows boxplot graphs for experiment # 1 for each instruction type. The times according to PAP are much longer than the other 4 which do not differ much from each other. According to Kruskal-Wallis the test is based on the p-value of the test $4.5*10^{-7}$. It is definitely less than the significance level of 5%, so we reject the null hypothesis and we can say that there is at least one group that significantly differs statistically from the others.
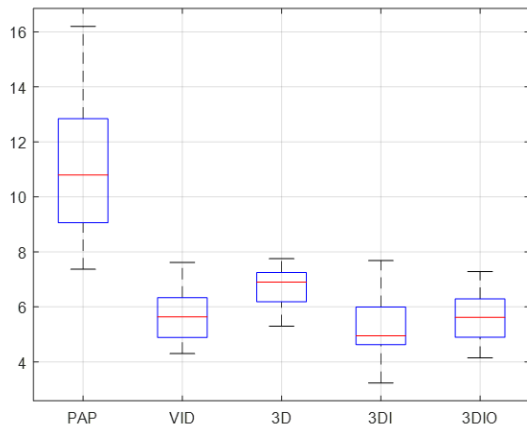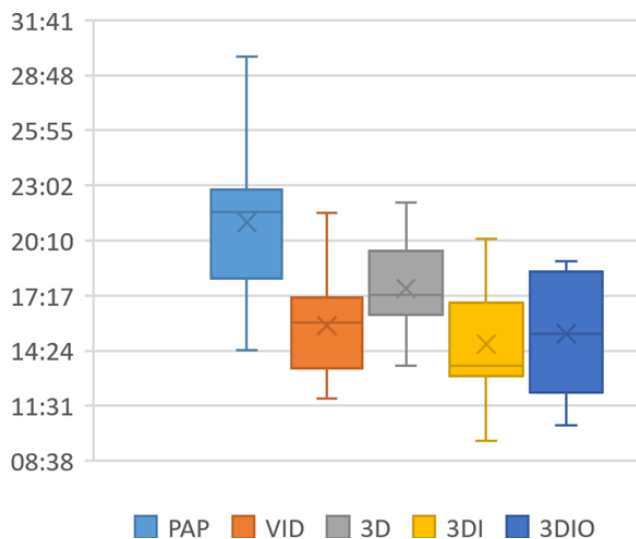
FIGURE 7. Box-plot for first try-out.

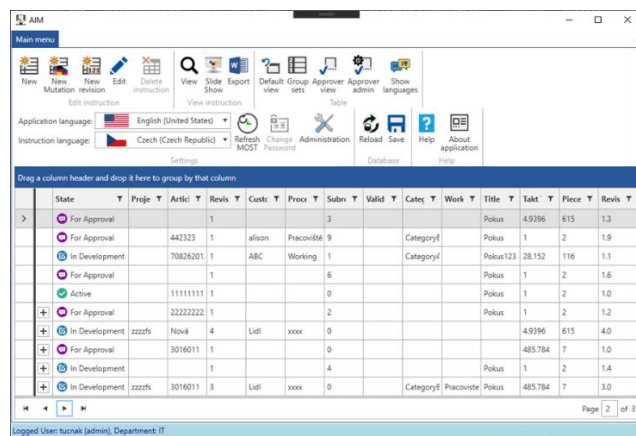**FIGURE 8.** Box-plot for all try-outs.

FIGURE 9. Assembly Instruction Management software – main menu screenshot.

If multiple comparisons are made, the P-value is less than 5% in three cases - i.e. there is a statistically significant difference between the PAP-VID, PAP-3DI, and PAP-3DIO

groups. This is confirmed by the rendered boxplots - 3D Boxplot is the PAP ''closest'', while the other three have shorter times. In general, however, these 4 do not differ significantly from each other, while the PAP guide differs from 3/4 of the rest. (It would differ from measurement 3 for a significance level of 10%). The statistical assessment of the other steps was similar. By way of illustration, a summary box-fence is shown for all the experiments.

In the statistical evaluation we did not deal with the correlation of height, weight, gender or skill.

At the end of the experiments, participants filled in a questionnaire with four questions to evaluate the tests and suggest improvements:

- Is this kind of learning beneficial? (scale)
- Do you think it could be technically improved –graphics, hardware, etc.? (scale)
- Describe any technical problems you encountered (open question)
- Suggest some improvements (open question)

The probands feel that this new way of learning is mostly beneficial, also the feedback on the technology was positive. The evaluation shows that the PAP group had a problem with orientation in the manual, which is confirmed by the measured times. No major shortcomings were found in the filmed instructions. This is due to the fact that when shooting the video, the hands were clearly visible during the process, which is not always possible. In the 3D group, the probands most often mentioned the problem of finding the right part, which was solved by the in-situ projection for the 3DI and 3DIO groups. In addition, also the problem with animation control was resolved for the 3DIO group. This problem was often mentioned in the 3D and 3DI groups.

VI. COMPARISON WITH OTHER STUDIES

Compared to other studies evaluating the effects of more modern solutions than paper, some show up to 30% better results, some show worse results (mostly in cases when using HMDs). Advantages of AR over paper by more than 32% were found in [32] (with significant error rate reduction), CAD animations vs paper by 37% in [3], similarly 20% better in [9] (although when assembling LEGO bricks). The VR assembly study [19] shows similar results. In [2], the error rate was half with instructions on a tablet, but these were still inferior to in-situ projected instructions – also when assembling LEGO. An opposing study showed practically opposite results [7].

The advantage of the 3DI and 3DIO groups over 3D can be confirmed by studies comparing picking tasks, where in-situ marking of correct containers make up for twice the performance compared to a normal paper list [20]. The picking operation can be considered as a part of an assembly operation.

It should be noted that our experiment was about on-site training. But we must also mention the option of off-site training, which is already used for training workers in

industry [42] or in other fields such as medicine [43]. We assume that the importance of off-site training will increase with the advent of technologies such as holographic displays [44] or precise hand tracking sensors [45].

VII. PRACTICAL IMPLEMENTATION

Whenever it is presented to manufacturing companies, so far production engineers have expressed excitement over the idea of replacing paper instructions with a more modern and more effective solution.

Employee turnover is causing considerable losses to manufacturing companies, and this is a way to reduce those losses. These companies have placed orders for first implementation so that they will be able to evaluate the results.

The system can produce interactive manuals – for example: The worker must confirm that they did not forget to install a seal or check whether a lever is moving correctly. So far there has been no demand for such functionality, all they wanted is a rendered video that would run in a loop on a screen near the workplace.

Among other things, there were requests to add icons with commands to wear protective gloves and glasses in the corner, displaying the overall layout of the workplace with a static character indicating where the operation takes place. The most difficult task was to animate the bending of a coolant distributor in air conditioners – which is a cluster of many pipes that need to be bent to reach their destination and to fit in a confined area, and they also must not interfere with each other.

There were also requirements that did not meet the system's ability to quickly and easily create the instructions. These were for example upholstery operations, although they still can be filmed and added instead of an animated step, disabling the possibility for AR in the step.

The main requirement of such implementation is a big enough screen capable of playing video or running a computer game. If the instructions are to be managed and updated automatically, a network connection is also needed to download current data from a local server. This can be a challenge for production halls that have been in the same place for a longer time; in new ones, this is usually not a concern as they are being prepared for implementation of Internet of Things or Industry 4.0.

Although this article focuses on the initial phase of understanding the manual (assembly instructions), it can be expected that, even during continuous use, it will have a positive impact not only on the production cycle, but also on reducing error rates through better understanding. We can also assume a reduction in the needed necessary skills of employees by suggested approach using (now it is often only a repetition of predefined movements).

Each particular implementation must be considered individually, as the technologies available and even the philosophies of the manufacturing companies differ and manuals would have to follow them.

Our approach of using Unity3D as the editor for assembly instructions supplemented with our own editor scripts proved to be feasible for non-expert users to learn. From our experience, we can teach 5 people how to use it in a 3 hour lesson. Only one of them found the work very difficult, although he was not very confident in computers overall. The others found the workflow easy but complained about making occasional mistakes. Their feedback was vital for further improvements.

The developed solution can be adapted and used for other aspects of the Smart City; it can be useful in hospitals, schools, public transport driver training, etc.

We have also developed a huge database system for Assembly Instruction Management (AIM). The instructions can be edited (with Word-like possibilities), quickly viewed as a slide-show and approved using defined patterns. AIM includes multilingual support. This information system is connected to our previously described software for virtual assemblies – different 3D manuals can be started in each step. This information software can be adapted to a larger scale in a big-data driven database (as shown in the ''extreme'' data-driven society of China modeled in [30]).

VIII. FUTURE PROGRESS

There are still some factors that need to be evaluated. So far the collected experience from practical implementations are positive, but new questions are emerging. These are mainly on how to present the information optimally, so that everybody (or most people at least) understand the presented task immediately or as fast as possible. [1] suggests using standardized words from simplified English, but icons would be more suitable. An ideal state would be to have widely accepted standardized and clearly understandable icons representing elementary tasks, like insert, screw, turn, etc. As the workers are international, the language in such manuals needs to be international as well.

Currently, the industry is awaiting the arrival of an AR capable device that would be able to meet the industrial requirements (mentioned in the introduction). When such a device is available, further testing needs to be carried out. As stated in [14], using instructions in the field of view (applicable also for AR) when conducting an assembly task when the worker already knows the process can yield worse results. Also, the physical and psychological effects of long-term use of such technologies remain unknown.

Further improvements in reducing non value-added tasks can be accomplished by controlling the instructions steps by other means than pressing a button. A waving gesture or voice command could be used. Even better, computer algorithms could be used to check the completion of the task and highlight possible errors.

As for our software implementations, we need to extend the portfolio of elementary and more complex animation templates and their editors so that they can be created as conveniently as possible.

IX. CONCLUSION

We have designed and tested a newly developed large scale extension for Unity3D capable of creating 3D animated assembly instructions. This software bundle was tested using non-expert computer programmers and graphics persons, who were, in most cases, able to learn to create such manuals in a short time. We have also tested the resulting assembly manuals against filmed and conventional paper manuals and yielded similar results as other studies, although not all measurements were significant according to the Kruskal-Wallis test. A water trap was used as the target assembly object for our experiment and the workplace was designed to be realistic and ergonomically optimized, to be as similar as possible to workplaces found in real production systems.

In answer to the question, what improvements have been made with the use of other variants compared to the paper guide? It can be stated that in the first attempt there was an improvement of 5:27 min (PAP - 11:18 min, Average others - 5:51 min). This means an improvement of 5:27 min (48% improvement). In the last (sixth) attempt, there was an improvement of 0:50 min (PAP - 3:24 min, Average other - 2:34 min), i.e. 25% improvement.

During implementation in real production systems, we have gained valuable experience and feedback that will drive both our research and development further. In this study we did not test Augmented Reality hands-free see-through helmet projected assembly manuals (we tested only AR In-Situ projection) as we are skeptical about the possibilities of deploying current AR solutions as well as [38], although we have prepared our software solution to be compatible with AR. The wearable AR is still problematic when we consider an assembly line with the work in short takt times (often in continuous operation with multiple shifts). In the industrial environment, the AR headsets should be targeted to those workers in the future. Current state AR headsets are not sufficiently easy-to-use for those kinds of applications. The work is monotonous and performed by less qualified workers. Using a current state AR headset on a whole day basis could result in possible health issues, described in [33] or [34]. Additionally, it is a risk for an employer to let a worker perform their job on such (still) expensive hardware. We know from our experiments with the high-end AR headset MS Hololens, that there are still other limitations such as – low brightness level, small field of view, latency, problems with tracking large-scale objects and the speed of object recognition mentioned in [41]. These and other shortcomings are described also e.g. in [35] and [36]. Therefore, we currently recommend using AR rather for maintenance as described e.g. in [39] or [40]. After overcoming the hardware issues, AR smart glasses will be a promising candidate as an everyday tool in a future Smart factory. We expect that our future research will include implementation of next-generation AR headsets for assembly instructions and comparison with this study and also probably with the previous one. An AR marker system with the same product and workplace but with different self-developed software was tested in our previous study [26] which shows that probands were 50% quicker after 10 tryouts.

The other significant future goal is to integrate this solution with other relevant Smart city branches.

REFERENCES

[1] M. Gattullo, A. E. Uva, M. Fiorentino, G. W. Scurati, and F. Ferrise, "From paper manual to AR manual: Do we still need text?" *Procedia Manuf.*, vol. 11, pp. 1303–1310, 2017.

[2] M. Funk, T. Kosch, and A. Schmidt, "Interactive worker assistance: Comparing the effects of *in-situ* projection, head-mounted displays, tablet, and paper instructions," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput. (UbiComp)*, 2016, pp. 934–939, doi: 10.1145/2971648.2971706.

[3] G. Watson, R. Curran, J. Butterfield, and A. C. Craig, "The effect of using animated work instructions over text and static graphics when performing a small scale engineering assembly," in *Proc. Collaborative Product Service Life Cycle Manage. Sustain. World*, 2008, pp. 541–550.

[4] S. Henderson and S. Feiner, "Exploring the benefits of augmented reality documentation for maintenance and repair," *IEEE Trans. Vis. Comput. Graphics*, vol. 17, no. 10, pp. 1355–1368, Oct. 2011.

[5] C. M. Robertson, B. MacIntyre, and B. N. Walker, "An evaluation of graphical context when the graphics are outside of the task area," in *Proc. 7th IEEE/ACM Int. Symp. Mixed Augmented Reality*, Sep. 2008, pp. 73–76.

[6] J. Tumler, F. Doil, R. Mecke, G. Paul, M. Schenk, E. A. Pfister, A. Huckauf, I. Bockelmann, and A. Roggentin, "Mobile augmented reality in industrial applications: Approaches for solution of user-related issues," in *Proc. 7th IEEE/ACM Int. Symp. Mixed Augmented Reality*, Sep. 2008, pp. 87–90.

[7] J. Blattgerste, B. Strenge, P. Renner, T. Pfeiffer, and K. Essig, "Comparing conventional and augmented reality instructions for manual assembly tasks," in *Proc. 10th Int. Conf. Pervas. Technol. Rel. Assistive Environments*, Jun. 2017, pp. 75–87.

[8] B. Schwerdtfeger, R. Reif, W. A. Gunthner, G. Klinker, D. Hamacher, L. Schega, I. Bockelmann, F. Doil, and J. Tumler, "Pick-by-vision: A first stress test," in *Proc. 8th IEEE Int. Symp. Mixed Augmented Reality*, Oct. 2009, pp. 115–124.

[9] A. Tang, C. Owen, F. Biocca, and W. Mou, "Comparative effectiveness of augmented reality in object assembly," in *Proc. Conf. Hum. Factors Comput. Syst.*, New York, NY, USA, 2003, pp. 73–80.

[10] D. W. F. Van Krevelen and R. Poelman, "A survey of augmented reality technologies, applications and limitations," *Int. J. Virtual Reality*, vol. 9, no. 2, pp. 1–20, 2010.

[11] D. Reiners, D. Stricker, G. Klinker, S. Müller, and A. F. Igd, "Augmented reality for construction tasks: Doorlock assembly," in *Proc. Int. Workshop Augmented Reality, Placing Artif. Objects Real Scenes, Placing Artif. Objects Real Scenes*, 1999, pp. 31–46, doi: 10.5555/322690.322694.

[12] D. Aiteanu, B. Hillers, and A. Graser, "A step forward in manual welding: Demonstration of augmented reality helmet," in *Proc. 2nd IEEE ACM Int. Symp. Mixed Augmented Reality*, 2003. pp. 309–310.

[13] M. Funk, A. Bächler, L. Bächler, T. Kosch, T. Heidenreich, and A. A. Schmidt, "Working with augmented reality?: A long-term analysis of *in-situ* instructions at the assembly workplace," in *Proc. ACM Int. Conf. Pervasive Technol. Rel. Assistive Environ.*, 2017, pp. 222–229, doi: 10.1145/3056540.3056548.

[14] M. Funk, T. Dingler, J. Cooper, and A. Schmidt, "Stop helping me–I'm bored!: Why assembly assistance needs to be adaptive," in *Proc. ACM Int. Joint Conf. Pervas. Ubiquitous Comput. Proc. ACM Int. Symp. Wearable Comput.*, New York, NY, USA, 2015, pp. 1269–1273.

[15] B. Schwerdtfeger, D. Pustka, A. Hofhauser, and G. Klinker, "Using laser projectors for augmented reality," in *Proc. ACM Symp. Virtual Reality Softw. Technol. (VRST)*, New York, NY, USA, 2008, pp. 134–137.

[16] N. Navab, "Developing killer apps for industrial augmented reality," *IEEE Comput. Graph. Appl.*, vol. 24, no. 3, pp. 16–20, May 2004.

[17] P. Fite-Georgel, "Is there a reality in industrial augmented reality?" in *Proc. 10th IEEE Int. Symp. Mixed Augmented Reality*, Oct. 2011, pp. 201–210.

[18] C. Knopfle, J. Weidenhausen, L. Chauvigne, and I. Stock, "Template based authoring for AR based service scenarios," in *Proc. IEEE Virtual Reality*, Dec. 2005, pp. 237–240.

[19] Z. Gibson, M. McGeown, J. Butterfield, A. Marzano, and A. B. Murphy, "Virtual reality as a tool to assist learning in aerospace assembly," presented at the Int. Manuf. Conf., Limerick, Ireland, Aug./Sep. 2016. [Online]. Available: https://pure.qub.ac.uk/en/publications/virtual-reality-as-a-tool-to-assist-learning-in-aerospace-assembl

[20] M. Funk, A. S. Shirazi, S. Mayer, L. Lischke, and A. Schmidt, "Pick from here!: An interactive mobile cart using *in-situ* projection for order picking," in *Proc. ACM Int. Joint Conf. Pervas. Ubiquitous Comput.*, 2015, pp. 601–609.

[21] M. Holm, O. Danielsson, A. Syberfeldt, P. Moore, and L. Wang, "Adaptive instructions to novice shop-floor operators using augmented reality," *J. Ind. Prod. Eng.*, vol. 34, no. 5, pp. 362–374, Jul. 2017, doi: 10.1080/21681015.2017.1320592.

[22] V. Alcácer and V. Cruz-Machado, "Scanning the industry 4.0: A literature review on technologies for manufacturing systems," *Eng. Sci. Technol., Int. J.*, vol. 22, no. 3, pp. 899–919, Jun. 2019.

[23] S. Werrlich, K. Nitsche, and G. Notni, "Demand analysis for an augmented reality based assembly training," in *Proc. 10th Int. Conf. Pervas. Technol. Rel. Assistive Environ.*, Island of Rhodes, Greece, Jun. 2017, pp. 416–422, doi: 10.1145/3056540.3076190.

[24] Eurostat. *Employment Statistics*. Accessed: Jul. 15, 2019. [Online]. Available: https://ec.europa.eu/eurostat/statistics-explained/index.php/Employment_statistics

[25] M. Richardson, G. Jones, and M. Torrance, "Identifying the task variables that influence perceived object assembly complexity," *Ergonomics*, vol. 47, no. 9, pp. 945–964, Jul. 2004, doi: 10.1080/00140130410001686339.

[26] P. Hoěí, "Augmented reality system for virtual training of parts assembly," *Procedia Eng.*, vol. 100, pp. 699–706, 2015, doi: 10.1016/j.proeng.2015.01.422.

[27] S. Marvin, A. While, M. Kovacic, A. Lockhart, and R. Macrorie, "Urban robotics and automation: Critical challenges, international experiments and transferable lessons for the UK," Univ. Sheffield, Sheffield, U.K., UK-RAS Netw. White Papers, 2018, doi: 10.13140/RG.2.2.16976.56321.

[28] L. V. Ivanenko, E. A. Karaseva, and E. P. Solodova, "Clusters, digital economy and smart city," in *Digital Transformation of the Economy: Challenges, Trends and New Opportunities*. Cham, Switzerland: Springer, 2020, pp. 291–295, doi: 10.1007/978-3-030-11367-4_28.

[29] P. Poulova, M. Cerna, J. Hamtilova, F. Malý, T. Kozel, P. Kriz, J. Han, and Z. Ulrych, "Hotel process simulator," in *Blended Learning: Educational Innovation for Personalized Learning*. Cham, Switzerland: Springer, 2019, pp. 128–136, doi: 10.1007/978-3-030-21562-0_11.

[30] F. Liang, V. Das, N. Kostyuk, and M. M. Hussain, "Constructing a data-driven society: China's social credit system as a state surveillance infrastructure," *Policy Internet*, vol. 10, no. 4, pp. 415–453, Oct. 2018.

[31] A. Visvizi and M. D. Lytras, *Smart Cities: Issues and Challenges : Mapping Political, Social and Economic Risks and Threats*. Cambridge, MA, USA: Elsevier, 2019.

[32] Z.-H. Lai, W. Tao, M. C. Leu, and Z. Yin, "Smart augmented reality instructional system for mechanical assembly towards worker-centered intelligent manufacturing," *J. Manuf. Syst.*, vol. 55, pp. 69–81, Apr. 2020, doi: 10.1016/j.jmsy.2020.02.010.

[33] I. Fajnerová, "Could prolonged usage of GPS navigation implemented in augmented reality smart glasses affect hippocampal functional connectivity?" *BioMed Res. Int.*, vol. 2018, pp. 1–10, Jun. 2018, doi: 10.1155/2018/2716134.

[34] B. Stock, T. P. dos Santos Ferreira, and C.-P. H. Ernst, "Does perceived health risk influence smartglasses usage?" in *The Drivers Wearable Device Usage*. Cham, Switzerland: Springer, 2016, pp. 13–23.

[35] P. Hübner, K. Clintworth, Q. Liu, M. Weinmann, and S. Wursthorn, "Evaluation of HoloLens tracking and depth sensing for indoor mapping applications," *Sensors*, vol. 20, no. 4, p. 1021, Feb. 2020, doi: 10.3390/s20041021.

[36] V. García-Vázquez, F. von Haxthausen, S. Jäckle, C. Schumann, I. Kuhlemann, J. Bouchagiar, A.-C. Höfer, F. Matysiak, G. Hüttmann, J. P. Goltz, M. Kleemann, F. Ernst, and M. Horn, "Navigation and visualisation with HoloLens in endovascular aortic repair," *Innov. Surgical Sci.*, vol. 3, no. 3, pp. 167–177, Oct. 2018, doi: 10.1515/iss-2018-2001.

[37] D. Jo and G. J. Kim, "*in-situ* AR manuals for IoT appliances," in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*, Jan. 2016, pp. 409–410, doi: 10.1109/ICCE.2016.7430669.

[38] M. Kerin and D. T. Pham, "A review of emerging industry 4.0 technologies in remanufacturing," *J. Cleaner Prod.*, vol. 237, Nov. 2019, Art. no. 117805, doi: 10.1016/j.jclepro.2019.117805.

[39] M. Gattullo, G. W. Scurati, M. Fiorentino, A. E. Uva, F. Ferrise, and M. Bordegoni, "Towards augmented reality manuals for industry 4.0: A methodology," *Robot. Comput.-Integr. Manuf.*, vol. 56, pp. 276–286, Apr. 2019, doi: 10.1016/j.rcim.2018.10.001.

[40] F. De Pace, F. Manuri, A. Sanna, and D. Zappia, "A comparison between two different approaches for a collaborative mixed-virtual environment in industrial maintenance," *Frontiers Robotics AI*, vol. 6, p. 18, Mar. 2019, doi: 10.3389/frobt.2019.00018.

[41] K. Židek, P. Lazorík, J. Piteä, and A. Hoáovský, "An automated training of deep learning networks by 3D virtual models for object recognition," *Symmetry*, vol. 11, no. 4, p. 496, Apr. 2019, doi: 10.3390/sym11040496.

[42] H. Zhang, "Head-mounted display-based intuitive virtual reality training system for the mining industry," *Int. J. Mining Sci. Technol.*, vol. 27, no. 4, pp. 717–722, Jul. 2017, doi: 10.1016/j.ijmst.2017.05.005.

[43] N. E. Seymour, A. G. Gallagher, S. A. Roman, M. K. O'Brien, V. K. Bansal, D. K. Andersen, and R. M. Satava, "Virtual reality training improves operating room performance," *Ann. Surgery*, vol. 236, no. 4, pp. 458–464, Oct. 2002, doi: 10.1097/00000658-200210000-00008.

[44] Z. He, X. Sui, G. Jin, and L. Cao, "Progress in virtual reality and augmented reality based on holographic display," *Appl. Opt.*, vol. 58, no. 5, p. A74, Feb. 2019, doi: 10.1364/AO.58.000A74.

[45] R. Jayasekera and X. Xu, "Assembly validation in virtual reality—A demonstrative case," *Int. J. Adv. Manuf. Technol.*, vol. 105, no. 9, pp. 3579–3592, Dec. 2019, doi: 10.1007/s00170-019-03795-y.

PETR HOŘEJŠÍ received the master's degree from the Faculty of Mechanical Engineering, University of West Bohemia, Pilsen, Czech Republic, with a specialization in industrial engineering, and the Ph.D. degree from the Department of Industrial Engineering and Management, University of West Bohemia. His Ph.D. dissertation is Parallel and Distributed Simulation Based on High-Level Architecture Remote Management Focused on Virtual Manufacturing Systems. Since 2006, he has been an Assistant Professor with the Department of Industrial Engineering and Management, University of West Bohemia. In 2019, he was an Assistant/Associate Professor (docent) in mechanical engineering. He completed his Habilitation work titled "Use of Virtual and Augmented Reality in Industrial Companies." His teaching areas include technical informatics; computer support in mechanical engineering; digital factory and virtual reality; database systems in CIM; practice in computer technology; methodological course (use of 3D data in archaeology); and virtual manufacturing systems of engineering enterprises (for Doctoral Study). His abroad experience includes a long-term stay on Universitá degli studi di Genova, Italy, and recurring teaching stays in Manchester Metropolitan University, GB, and Žilinská univerzita v Žiline, Slovakia. His projects focus on, for instance, on Risk Environment Simulator (RES)–VR simulator for alcohol addicts; a huge planning software; an IS for a personal agency; virtual classroom project; and various serious games.

KONSTANTIN NOVIKOV received the Bc. and Ing. (master's) degrees from the Department of Industrial Engineering and Management, Faculty of Mechanical Engineering, University of West Bohemia, Pilsen, Czech Republic, in 2015 and 2017, respectively, and the M.B.A. degree in financial management and accounting from the Cambridge Business School, Prague, in 2020. He is currently pursuing the Ph.D. degree with the Department of Industrial Engineering and Management, Faculty of Mechanical Engineering, University of West Bohemia. His bachelor's work includes the internal audit of industrial companies. His diploma work is costs evaluation of heating elements production. From 2016 to 2017, he was an Industrial Engineer in cost evaluation of heating elements production with Arbonia Forster Holding AG, Arbon, Switzerland. From 2014 to 2018, he was an Industrial Engineer at the Lean Production Department, Kermi s.r.o., Stříbro, Czech Republic. Since 2018, he was a Project Coordinator with CIE consulting s.r.o., Pilsen, where he plays a role as the Team Leader, with an aim of project management. He is also responsible for the project on the professional and project side, from the first meeting with the customer to project submission. His core area of the job lies in lean production, logistics, and virtual reality. His teaching areas include industrial engineering, methods of industrial engineering, and logistics. His abroad experiences include long-term stays at Chienkuo Technology University, Chienkuo, Taiwan, and Cukurova University, Adana, Turkey, as well as a short-term stay at the University of Maribor, Slovenia. His projects focus on employee training in virtual reality, including many implementations of virtual work instructions, and the design and optimization of workplaces in virtual reality.

MICHAL ŠIMON received the master's degree from the Department of Industrial Engineering and Management, Faculty of Mechanical Engineering, University of West Bohemia, Pilsen, Czech Republic, in 1999, and the Ph.D. degree from the Department of Industrial Engineering and Management, Faculty of Mechanical Engineering, University of West Bohemia, under a thesis "Knowledge system to support innovation of selected business processes." He was an Assistant/Associate Professor (docent) Habilitation in advanced methods for higher the production capacity of the production system from the Faculty of Mechanical Engineering, University of West Bohemia, in 2008. Since 2007, he has been an Assistant Professor with the Department of Industrial Engineering and Management, University of West Bohemia, where he has also been the Head of the Department since 2009. From 2014 to 2019, he was a Project Coordinator and Owner of CIE s.r.o., Pilsen, where he was the Team Leader with an aim of project management. His teachings include industrial engineering, production control, and logistics. His projects focus on the development of a new generation of short handguns with regard to human ergonomics and anthropometry; software for evaluating the long-term impact of workplace layout on worker health; and OptiPro 4.0: Optimization of production processes—Industry 4.0 im bayerisch-tschechischen Grenzraum.

• • •