

Objectively measuring learning outcomes of information technology-assisted training courses

A pilot study of
a novel
Performance
Indicator

437

Gerald Schneikart and Walter Mayrhofer

*Institute for Digital Transformation and Strategy, FH Wien der WKW,
Vienna, Austria*

Received 15 April 2022
Revised 21 August 2022
Accepted 21 August 2022

Abstract

Purpose – The objective of the presented pilot study was to test the applicability of a metric to specifically measure performance improvement via a hands-on workshop about collaborative robotics.

Design/methodology/approach – Candidates interested in acquiring basic practical skills in working with a collaborative robot completed a distance learning exercise in preparation for a hands-on training workshop. The candidates executed a test before and after the workshop for recording the parameters compiled in the tested performance index (PI).

Findings – The results reflected the potential of the tested PI for applications in detecting improvement in practical skill acquisition and revealed potential opportunities for integrating additional performance factors.

Research limitations/implications – The low number of candidates available limited in-depth analyses of the learning outcomes.

Practical implications – The study outcomes provide the basis for follow-up projects with larger cohorts of candidates and control groups in order to expedite the development of technology-assisted performance measurements.

Social implications – The study contributes to research on performance improvement and prediction of learning outcomes, which is imperative to this emerging field in learning analytics.

Originality/value – The development of the presented PI addresses a scientific gap in learning analytics, i.e. the objective measurement of performance improvement and prediction along skill-intensive training courses. This paper presents an improved version of the PI, which was published at the 12th Conference on Learning Factories, Singapore, April 2022.

Keywords Distance learning, E-Learning, Hybrid learning, Learning analytics, Learning efficacy, Learning outcome, Performance development, Performance index, Performance measurement, Complex skills acquisition

Paper type Research paper

1. Introduction

Delivering skill-intensive knowledge is often done by intermittently providing academic preparation and practical application of the theoretical input. It is a commonly held belief that alternating “studying” and “application” sequences leads to a steeper learning curve and better knowledge-retention rates. However, there are few canonical methodologies or metrics

© Gerald Schneikart and Walter Mayrhofer. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licences/by/4.0/legalcode>

The authors would like to thank Tanja Zigart and Patrick Killingseder for their kind assistance in establishing the test procedure, Clara Fischer, Maximilian Papa and Sebastian Schlund for the cooperation in the EIT Manufacturing funded research project “Strengthening skills and training expertise in human machine interaction (ENHANCE)” by TU Wien, that produced the first version of the Performance Indicator presented in this paper and Ann Ilaria Mayrhofer for editing this article.



to objectively demonstrate that one practical learning course is superior over another. The most objective means for measuring learning effectiveness are still in the forms of quizzes, score-based evaluations or similar approaches (Bonde *et al.*, 2014; Makransky *et al.*, 2016a, b), which can be automatically evaluated by applying algorithms. Nevertheless, these forms of evaluation are more applicable for measuring explicit, factual knowledge since they directly scan an individual's memory (Curran, 1997).

In case of field-tests, performance improvements are usually evaluated by tutors or teachers (Bach *et al.*, 2014) in a summative manner. In the context of the *New World Kirkpatrick Model*, those means of evaluation represent "required drivers" of the Level 3 (see Figure 1), which are "processes and systems that reinforce, monitor, encourage, and reward performance of critical behaviors on the job" (Kirkpatrick *et al.*, 2016). However, evaluation by tutors or teachers is still a highly subjective approach and when it comes to evaluation of an individual's learning progress of a complex task – the "know-how-to-do" – there is little methodology. In other words, there is a lack of a comprehensive metric in learning analytics that enables an objective measurement of performance improvement rates along complex paths of practical training in various disciplines (e.g. collaborative robotics).

In 1959, the researcher Donald Kirkpatrick published a four-level model for evaluating training programs (Figure 1) (Kirkpatrick, 1959a, b, 1960a, b).

In 2016, Jim Kirkpatrick and Wendy Kirkpatrick, the oldest son and daughter-in-law of Don Kirkpatrick, published an updated version of his original four-level model for evaluating training programs (Kirkpatrick, 1959a, b, 1960a, b, Kirkpatrick and Kirkpatrick, 2006; Kirkpatrick *et al.*, 2016). According to the *New World Kirkpatrick Model*, Levels 1 and 2 of the original four-level model became over-emphasized over the decades since its publication in the late 1950s, because the Levels 3 and 4 were "too expensive and difficult to evaluate" (Kirkpatrick *et al.*, 2016). Digitalization of business and learning processes in particular turned the application of all four levels into a challenge. The *New World Kirkpatrick Model* adds new elements to the unfairly neglected Level 3, stating "The degree to which participants apply what they learned during training when they are back on the job", and Level 4: "The degree to which targeted outcomes occur as a result of the training and the support and accountability package". The new Level 4 was complemented with "Leading Indicators" that help interpreting critical behaviors. This should bring the desired results as those "critical behaviors" are part of the new Level 3, which were supplemented with "required drivers" and "on-the-job learning" (Kirkpatrick *et al.*, 2016).

In order to address the problem of objectively measuring the performance improvement rate for practical knowledge development, a prototypic performance index (PI) has been derived. It is aligned with a similar index previously published (Guneysu Ozgur *et al.*, 2020), albeit specifically optimized for working with collaborative robots (cobots) and with the outlook to further develop by integrating additional factors and weighed parameters. The prototypic PI presented in this paper compounds the time required to perform a task (working

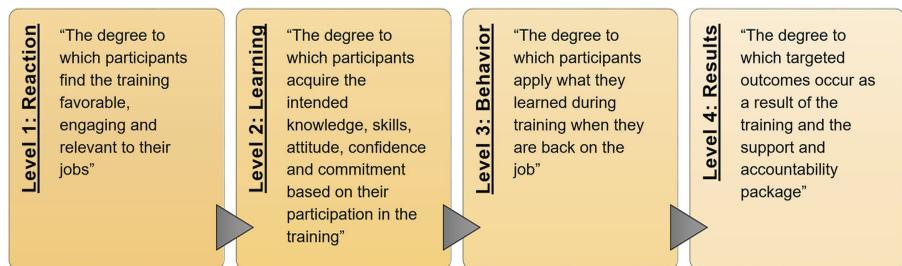


Figure 1.
The four levels of
Kirkpatrick's original
evaluation model

Source(s): Own figure derived from (Kirkpatrick, 1959/60; Kirkpatrick *et al.*, 2016)

time), the extent of needed assistance and error rates while performing the task. In addition, this paper showcases results of a pilot study on the PI applied to a combined distance/hands-on practical hybrid learning environment to build knowledge and skills in collaborative robotics. After having demonstrated the PI's potential for applications as a prediction tool for learning outcomes (Mayrhofer *et al.*, 2022), the objective of this study was to test an improved version of the PI in order to generate first indicative results, which point toward including additional factors to be considered for an improved performance metric.

The distance-learning component was implemented with a training module on the *Skills.move* platform by EIT-Manufacturing, covering the overall topic of working and implementing cobots. *Skills.move* is a newly-established learning platform that offers flexible learning experiences. It is the implementation of the Guided Learning Platform that was presented at the 11th Conference on Learning Factories (Mayrhofer *et al.*, 2021). In a nutshell, *Skills.move* offers training courses in the area of manufacturing (i.e. additive manufacturing, robotics, cobotics, automation, etc.) organized as learning paths comprising short, consecutive training sessions (using slide decks, video clips, documentations, or quizzes), and were dubbed “learning nuggets”. The long-term vision is to provide learners with the option to adjust the learning intensity based on their individual needs, talents, experience and learning style. To date, this “adaptive learning” approach has only been realized in fragments. However, the ability to complete the learning nuggets on-demand, when learners feel ready, and the possibility to change the sequence of the nuggets or repeat certain content improves the engagement of learners (*Skills.move* users). While this teaching strategy seems efficient in terms of user satisfaction, a prerequisite for the users’ learning motivation (Heckman, 2008; Kyllonen *et al.*, 2014; Levin, 2012), the study indicated that a combination of improved interactive and hands-on content would be desirable (Mayrhofer *et al.*, 2021). Therefore, *Skills.move* might be ideally applied as a preparation tool for follow-up practical training sessions, where users have the opportunity to complement their explicit knowledge acquired on *Skills.move* (hard skills) with tacit or implicit knowledge (development of soft skills), which is fundamental to performance improvement (Curran, 1997).

2. Methods

Potential candidates with various backgrounds and professions, but no education in robotics were invited to participate in the pilot study. Candidates that accepted the invitation received a participation sheet with details about the assignment to the distance-learning courses and the workshop approximately one week before the workshop commenced. In order to participate, interested candidates had to give their informed consent and subsequently were enrolled to the training module ME1 - Module Executives 1 offered on the EIT's Guided Learning platform *Skills.move* (<https://www.skillsmove.eu>). The learning paths comprised learning nuggets of five to 10 min designed to build basic knowledge about cobotics.

Five candidates participated in the pilot study. In the week leading up to the workshop days, the participants were asked to complete the learning nuggets in preparation for an ensuing hands-on training in working with the educational cobot *e.Do*. by COMAU S.p.A. In the context of a workshop organized at the pilot factory of TU Wien, participants received tutoring in interacting with *e.Do* and learned how to program the cobot in order to automatize a “pick-and-place” procedure with a rubber ball.

After a comprehensive introduction to mandatory safety rules in working with cobots, participants were asked to take a practical test specifically chosen to evaluate the candidates' individual performance in working with *e.Do*. The assigned task was to program a “pick-and-place” use-case procedure that enabled *e.Do* to automatically pick up a rubber ball from a cylinder and place it into another. The programming of the entire “pick-and-place” process can be described as a sequence of ten working steps: (1) create a project, (2) lower the robot

arm to bring the gripper in close proximity to the ball, (3) bring the gripper into an exact position to grasp the ball, (4) grasp the ball, (5) lift the ball, (6) re-position the robot arm with the ball to the second cylinder, (7) lower the ball onto the second cylinder, (8) release the ball, (9) lift the robot arm and (10) bring the robot arm back to its starting position.

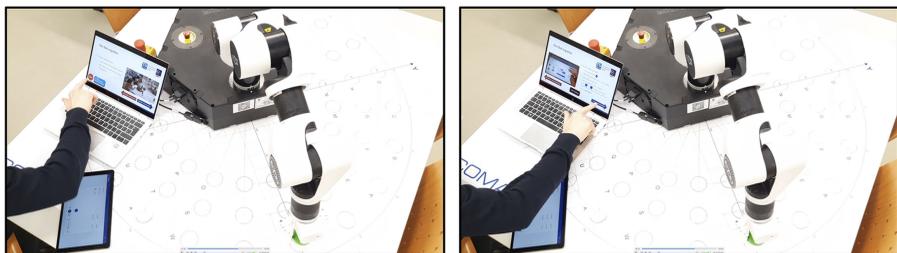
In order to interact with *e.Do*, candidates were provided with a tablet running the *e.Do* app. Candidates were free to find an optimal solution for the task themselves, although they were also offered guidance comparable to a users' manual. The screenshots in Figure 2a show the manual's user interface accessible with the touchscreen of a laptop. On the start screen, a short task description and a video displayed the desired outcome of the test. After pushing an interactive arrow button, the first working step was described in a short text and an illustrative video clip. Candidates requiring further assistance had the option to tap an interactive "HELP" button in the lower left corner of the screen that opened a detailed operating manual in addition to a video clip that demonstrated how to work through the current step. Arrow buttons at the bottom of the screen allowed navigation between consecutive working steps, while a list of all working steps leading to an optimal solution was also accessible from each screen.

In order to measure the parameters required to assess the candidates' performance, a camera capturing the work area recorded the entire test activities.

Working times (measured in minutes) for individual working steps were recorded, starting from the moment the candidates tapped the interactive arrow button opening the description



(a)



(b)

Figure 2.
Pre- and post-test
(a) User interface
(b) Snapshot of a candidate consulting assistance (left) and switching to a next working step during the pre-test (right)

window of the current working step and ending when they switched to the next one. If the candidate worked without instruction (which was often the case in the post-test phase), start and end time points of an individual working step were the initial and last tap to move and program the robot for that specific working step respectively.

The times for those working with detailed assistance were registered starting from the moment the candidates tapped the red “HELP” button (Figure 2b, left) until the moment, when they switched to the next working step (Figure 2b, right).

After completing the working steps, the “pick-and-place” sequences programmed by each candidate were executed. Incomplete or undesired movements of the robot arm (e.g. wrong or omitted movement directions) were considered to be mistakes made when programming the respective working step.

The test before the actual workshop started (pre-test) was the same as the final test after the workshop had ended (post-test).

In order to estimate the learning efficacy of the practical workshop, we proposed a comprehensive PI that combines different performance criteria. The candidates’ individual PIs at the pre- and post-tests were calculated according to a formula derived from a similar index previously published (Guneysu Ozgur *et al.*, 2020). However, the PI of the current study was adapted by incorporating performance parameters relevant to the “pick-and-place” task, including the working time to complete the entire task ($T_{\text{candidate}}$), working time when assistance was needed ($T_{\text{assistance}}$) and mistakes in the programmed sequence, which all negatively correlate with performance. In order to account for mistakes (erroneously programmed working steps resulting in unintended movements of the robot arm) to be detected by the PI, a time penalty per mistake (T_{mistakes}) was introduced as a placeholder.

$$\text{PI} = \frac{1}{T_{\text{candidate}} + T_{\text{assistance}} + T_{\text{mistake}}} \quad (1)$$

The analyses of the individual video footage recorded during the pre- and post-test allowed for the measurement of the single performance parameters.

The results of the pre- and post-test were analyzed with the *R* software environment using the packages “tidyverse” (<https://cran.r-project.org/package=tidyverse>) and “fmsb” (<https://cran.r-project.org/package=fmsb>).

3. Results

3.1 The candidates worked faster at the post-test

The radar charts in Figure 3 show that all candidates spent the most time programming *e.Do* to lower the gripper arm toward the rubber ball, suggesting this was the most demanding working step in terms of skill and thus had a stronger impact on the PI than the other working steps.

Candidate #1 needed more than nine minutes, whereas candidate #4 managed this working step in approximately three minutes during the pre-test; in comparison, the control person, a professional engineer, finished this part in around one minute. Candidate #5 struggled the most with positioning the gripper to grasp the ball and eventually gave up, which is the reason for the 0 minutes-records of the residual working steps.

After participating in the workshop, all candidates improved their performance with respect to overall working time, except candidate #4 who required an additional minute, compared to the pre-test. The candidates’ improvements result mostly from shorter working times needed to lower the gripper arm, supporting the view that this working step was the most demanding. Candidates #1, #2, #3 and #5 programmed this working step by up to four minutes quicker compared to the pre-test; candidate #4 worked more than one minute longer on this step, which explains the prolonged total working time to solve the entire programming task. Even though candidate #5 also improved, the candidate again gave up during the exact positioning of the gripper.

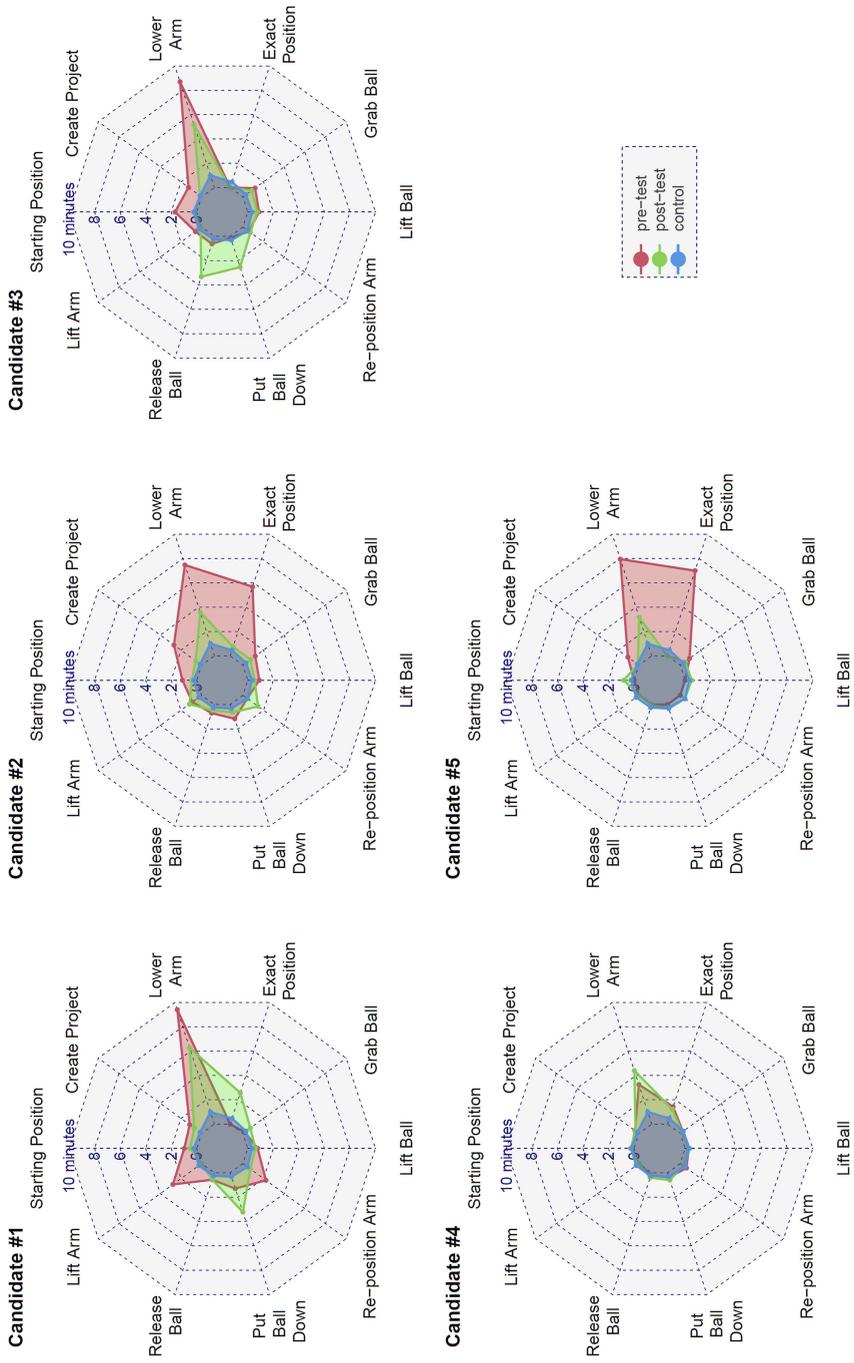


Figure 3.
Required working time by each candidate to accomplish the individual working steps

Interestingly, candidates #1 and #3 spent more than two minutes when placing the ball onto the second cylinder during the post-test. Presumably, both candidates underestimated this procedure or overestimated their perceived skills after having participated in the workshop, as neither of them sought out assistance for completing this maneuver during the post-test (Figure 4).

3.2 Most of the candidates worked without assistance at the post-test

The analysis of the working time records concerning when candidates worked with detailed assistance further support the view that the working step to lower and position the gripper for grasping the rubber ball was the most difficult part of the task. The radar charts in Figure 4 show that four candidates needed help during the pre-test; except candidate #4, who worked without assistance. After the workshop, all candidates were able to work without assistance at nearly all working steps; only candidate #1 needed help to perform working step 2.

3.3 Almost all candidates managed to establish functional “pick-and-place” sequences at the post-test

Running the programmed “pick-and-place” sequences of the individual candidates further indicated, unsurprisingly, that they improved in working with *e.Do* during the workshop. Before the workshop, only candidate #4 managed to establish an automatized sequence, which moved the rubber ball from one cylinder to the second one without dropping it. However, after the workshop, the majority of the sequences automatized during the post-test worked as intended, except the sequence by candidate #5, who stopped both tests after the working step of lowering the gripper arm. Nevertheless, candidate #5 still managed to establish at least one sequence during the post-test that enabled the cobot to grasp the ball (Figure 5).

The observation that only candidate #4 managed to deliver a working sequence at the pre-test (Figure 5) while finishing as the quickest (Figure 3) and without assistance (Figure 4) suggests that candidate #4 had already acquired experience in working with cobots before the study. When queried, candidate #4 confirmed acquiring previous experience with *e.Do*.

3.4 Derivation of the performance index (PI)

Overall, the analyses of the performance parameters already showed that the candidates improved their skills in applying the *e.Do* cobot for automatization purposes. However, the results also indicated that several additional factors needed to be considered when establishing a performance metric. The results clearly suggested that performance metrics should be capable of detecting incomplete work. For example, candidate #5 stopped both the pre- and post-tests after the working steps of lowering and grasping the ball, resulting in 0-min working times for the residual working steps (Figure 3). Since this should result in a lower PI development when compared with the other candidates, all of whom finished the task, an additional time-penalty factor “ $T_{\text{surrender}}$ ” was added to account for working steps that could not be recorded.

$$PI = \frac{1}{T_{\text{candidate}} + T_{\text{assistance}} + T_{\text{mistake}} + T_{\text{surrender}}} \quad (2)$$

The extent of time penalties per single mistake or per surrendered working step was set as 10 times the minutes needed by the control person to finish the n th working step ($T_{n[\text{control}]}$).

The outcome of the post-tests further suggested that the individual working steps have different influences on the success of the task. For example, some candidates skipped the step of bringing the gripper into an exact position to grasp the ball (resulting in zero working time

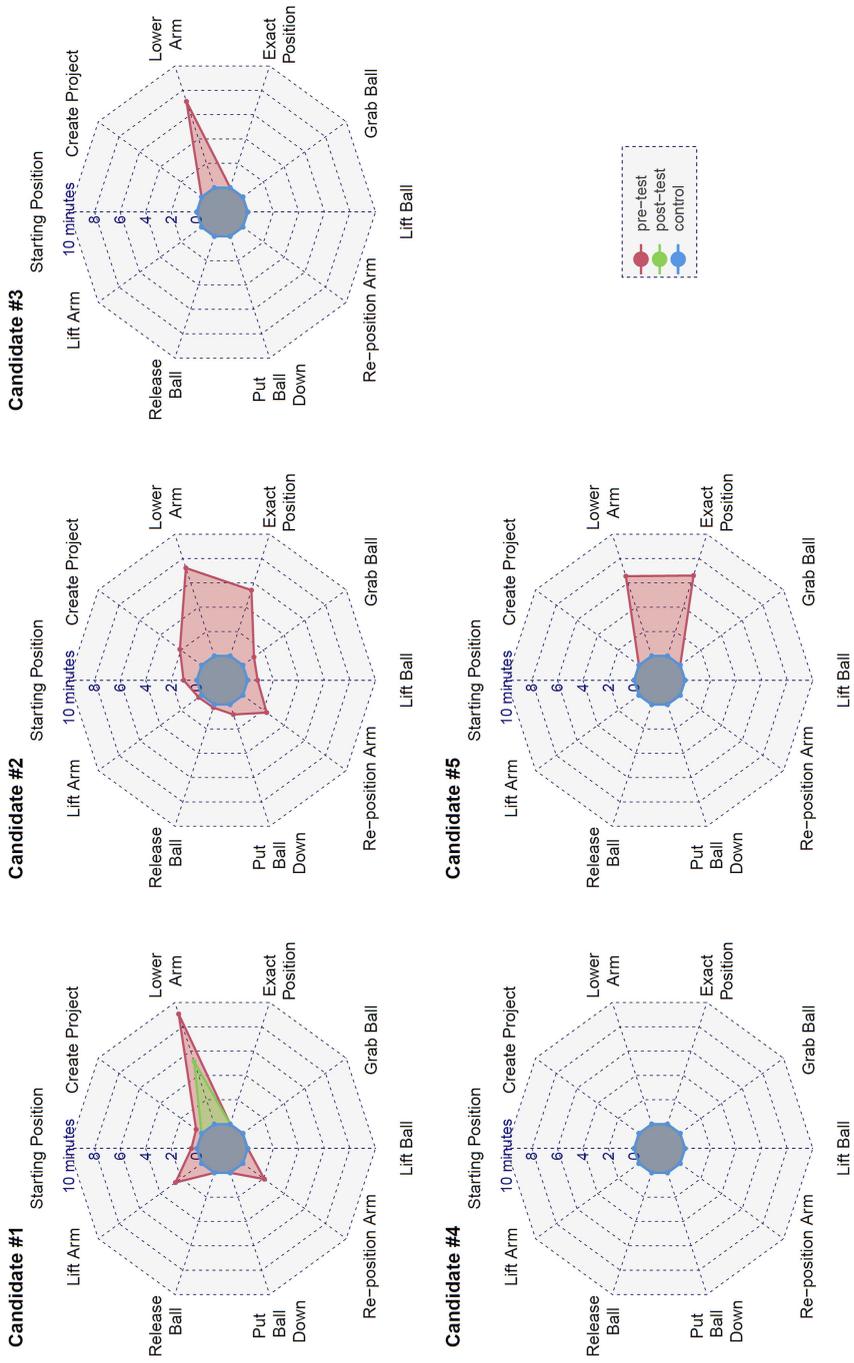


Figure 4. Working time with detailed assistance to accomplish the individual working steps

Pre-test

Working step	Candidate #1	Candidate #2	Candidate #3	Candidate #4	Candidate #5	Control
Create Project	0	0	0	0	1	0
Lower arm	0	0	0	0	1	0
Exact position	1	0	0	0	1	0
Grab ball	1	0	0	0	1	0
Lift ball	1	0	0	0	1	0
Reposition arm	0	0	0	0	1	0
Put ball down	1	1	1	0	1	0
Release ball	1	0	0	0	1	0
Lift arm	1	0	0	0	1	0
Starting position	0	0	0	0	1	0

Post-test

Working step	Candidate #1	Candidate #2	Candidate #3	Candidate #4	Candidate #5	Control
Create Project	0	0	0	0	0	0
Lower arm	0	0	0	0	0	0
Exact position	0	0	0	0	0	0
Grab ball	0	0	0	0	0	0
Lift ball	0	0	0	0	1	0
Reposition arm	0	0	0	0	0	0
Put ball down	0	0	0	0	1	0
Release ball	0	0	0	0	1	0
Lift arm	0	0	0	0	1	0
Starting position	0	0	0	0	1	0

Mistake	1
No mistake	0

Figure 5. Mistakes in individual working steps of candidates' "pick-and-place" sequences

of the working step "exact position" in Figure 3), but still managed to establish a functional "pick-and-place" sequence during the post-test (Figure 5). Completion of a task despite skipped or unfinished working steps clearly has a positive impact on performance. This suggests that the sum of weighted performance parameters corresponding to each working step needs to be applied in the PI.

$$PI = \frac{1}{\sum_{i=1}^n ((T_{n[\text{candidate}]} + T_{n[\text{assistance}]} + T_{n[\text{mistake}]} + T_{n[\text{surrender}]}) * w_n)} \quad (3)$$

$T_{n[\text{candidate}]}$... Time needed by the candidate to accomplish the n th working step

$T_{n[\text{assistance}]}$... Time working with assistance at the n th working step

$T_{n[\text{mistake}]}$... Time penalty of $10 \times T_{n[\text{control}]}$ for a mistake in the n th working step

$T_{n[\text{surrender}]}$... Time penalty of $10 \times T_{n[\text{control}]}$ for the missing n th working step after surrendering the task

w_n ... Relevance (weight) of the n th working step for the completion of the task

In the PI presented in this paper, the working time needed by the control person to accomplish the n th working step normalized to the time needed by the same control person to complete the entire task served as a weighing factor.

$$PI = \frac{1}{\sum_{i=1}^n \left((T_{n[\text{candidate}]} + T_{n[\text{assistance}]} + T_{n[\text{mistake}]} + T_{n[\text{surrender}]}) * \frac{T_{n[\text{control}]}}{T_{\text{total}[\text{control}]}} \right)} \quad (4)$$

$T_{n[\text{control}]}$... Time needed by the control person to accomplish the n th working step

$T_{\text{total}[\text{control}]}$... Total time needed by the control person to complete the entire task

In order to achieve a dimensionless PI, each performance parameter was normalized to the performance parameter of a professional tutor highly experienced in working with *e.Do* (control). **Formula (5)** was then applied to calculate the candidates' individual PI.

$$PI = \frac{1}{\sum_{i=1}^n \left(\left(\frac{T_{n[\text{candidate}]}}{T_{n[\text{control}]}} + \frac{T_{n[\text{assistance}]}}{T_{n[\text{control}]}} + \frac{T_{n[\text{mistake}]}}{T_{n[\text{control}]}} + \frac{T_{n[\text{surrender}]}}{T_{n[\text{control}]}} \right) * \frac{T_{n[\text{control}]}}{T_{\text{total}[\text{control}]}} \right)} \quad (5)$$

3.5 Almost all candidates improved working with *e.Do* at the workshop

Comparing the candidates' relative PIs suggests that the prototypic PI has high potential as a meaningful metric to measure performance improvement. The PI of candidates #1, #2, #3 and #5 improved from the pre- to the post-test timepoint. The performances of candidates #1 and #2 were comparable at the pre-test, but candidate #1, the only individual working with assistance at the post-test, revealed a flatter learning curve (**Figure 6**). The slope of the learning curve is representative of the learning efficacy. In this sense, candidate #2, who achieved the steepest learning curve, was the most responsive to the learning input provided by the workshop. As expected, candidate #5, who gave up at both the pre- and at the post-test, reached the flattest learning curve. Unsurprisingly, candidate #4, who demonstrated the best performance parameters, had the highest PI at the pre- and post-test.

4. Discussion

Candidate #4 was the only one with a drop in performance, which is in line with the additional minute required to bring the gripper arm into position to grasp the ball (**Figure 3**). Presumably, a factor not included in the tested PI was the candidate's concentration after an all-day workshop, an indication that further development of the presented prototypic PI might also integrate mental parameters. This could potentially be achieved with technologies

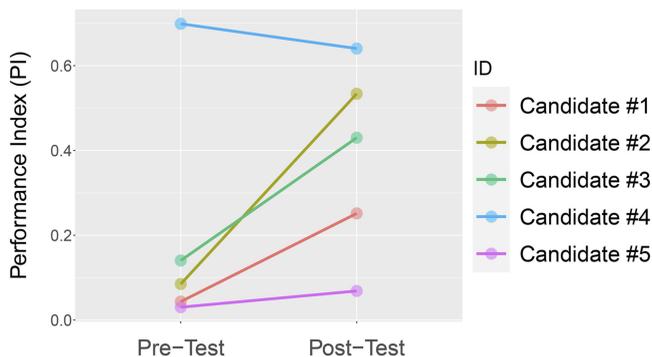


Figure 6.
Performance
improvement of
each candidate

enabling automated emotion recognition (Dzedzickis *et al.*, 2020), which would have the added advantage of preserving objectivity during the performance measurement.

Candidates #1, #2, #3 and #4 started with different PI levels varying in the range from 0.0305 to 0.1405. In addition, their individual performance improvements varied, as indicated by the different slopes of their learning curves. These results suggest that further developments of PIs should involve a person's talent or potential to improve specific skills. Since such factors are determined by one's attitude and willingness to learn (Heckman, 2008), (subjective) skill measurements pertaining to Levels 1 and 2 of the *New World Kirkpatrick Model* may also have the potential to detect baseline performances and could thus potentially be applied in advanced forms of PI metrics for the prediction of performance development (Kirkpatrick *et al.*, 2016).

Additionally, the analysis revealed that performance metrics derived from a sequence of individual working steps should take into account their relative difficulty levels, which was also concluded by Bach *et al.* (2014). For example, the working step of lowering the gripper arm seemed to be the most demanding one in terms of skill as candidates spent the most time on this step when compared to the others (Figures 3 and 4).

The performance measures derived from candidate #4 underpinned the power of the presented method for measuring performance. Candidate #4 showed the best performance by far as the candidate eventually confirmed.

Together, these observations demonstrate the PI's potential to detect or predict different learning outcomes of groups of candidates in studies on novel learning paths, learning modalities or technologies designed to support practical learning of skill-intensive, hands-on activities, e.g. the application of virtual reality to prepare students for wet-lab activities in a safe and cost-efficient manner (Schneikart, 2020). In addition, using an application as a prognostic tool in learning analytics to, for example, forecast learning outcomes or to support decision-making is highly conceivable (Leitner *et al.*, 2017; Siemens, 2013).

In this regard, it should be pointed out that applications of PI metrics to compare non-related tasks, learning methods, processes or technologies might not be feasible. As mentioned above, the results suggest that performance is a product of multiple factors, which are only indirectly measured by the presented approach. Additional factors that could influence performance are, for example, a person's mental state, interest to learn as well as pre-existing or complementary competencies.

Competency in particular plays a particular role in performance. Acquiring a new practically-oriented skills often starts with gathering data and information about the process and the environment where the specific skill is to be applied. According to the DIKW (Data-Information-Knowledge-Wisdom) pyramid, applying information results in knowledge and ultimately wisdom (Ackoff, 1989; Davenport and Prusak, 1998; Schon, 1983; Schon and DeSanctis, 1986). However, the DIKW-pyramid usually refers to the cognitive domain. With regard to physically-oriented abilities, an extension of the pyramid with a competence-component described as "practical wisdom" or "knowledge in action" seems necessary (Lalor *et al.*, 2015). In this respect, the term competence combines knowledge, skills and attitudes according to Roe's definition "a learned ability to adequately perform a task, duty or role" (Roe, 2002). Therefore, competency is a role-specific determinant defined by talent and skills (knowledge and experience) (Kuruba, 2019), which confines performance to a specific task or job. The PI presented in this paper is intended to provide a metric for the process of complex skills acquisition that lies at the core of competency development.

5. Conclusions

This paper presents an emerging approach to bridge a scientific gap in learning analytics, which is the objective measurement of performance development and prediction. This problem was addressed by developing a comprehensive PI that combines working

time to complete a task, time working with assistance and mistakes (deviations from the intended results). The objective of the conducted pilot study was to assess the capability of the developed PI to detect performance improvements of candidates volunteering to learn basic practical skills in working with cobots. Even though the analysis showed room for improvement, the pilot study demonstrated the PI's potential as a meaningful metric for performance improvement during skill-intensive training programs.

The outcome of this study thus provides the basis for future studies with larger cohorts of candidates and control groups, which will allow empirical–statistical analysis of learning outcomes, and the exploration of additional factors influencing an individual's performance. In particular, the consideration of competency will be of high interest as it plays an important role in performance. A method to measure competency has already been published (Kuruba, 2019), which will be considered for integration with the presented approach of PI measurement.

Since the manual data acquisition proved to be very tedious, further projects to explore the possibility of automatic performance measurements should be envisaged. In the ongoing era of digitalization, automatic detection of performance parameters will confront the development of performance metrics toward big data problems, which acquires increasing importance in learning analytics (Hadavand *et al.*, 2019). At the same time, this will create new applications for machine learning technologies, which could enable more precise performance measures. In fact, in the rapidly developing field of learning analytics (Wong *et al.*, 2018), first reports on automatic performance measures (Guitart *et al.*, 2015; Loh, 2013) and applications of machine learning are already available (Khosravi *et al.*, 2021).

References

- Ackoff, R.L. (1989), "From data to wisdom", *Journal of Applied Systems Analysis*, Vol. 16 No. 1, pp. 3-9.
- Bach, C., Miernik, A. and Schönthaler, M. (2014), "Training in robotics: the learning curve and contemporary concepts in training", *Arab Journal of Urology*, Vol. 12 No. 1, pp. 58-61, doi: [10.1016/j.aju.2013.10.005](https://doi.org/10.1016/j.aju.2013.10.005).
- Bonde, M.T., Makransky, G., Wandall, J., Larsen, M.V., Morsing, M., Jarmer, H. and Sommer, M.O.A. (2014), "Improving biotech education through gamified laboratory simulations", *Nature Biotechnology*, Vol. 32 No. 7, pp. 694-697, doi: [10.1038/nbt.2955](https://doi.org/10.1038/nbt.2955).
- Curran, T. (1997), "Effects of aging on implicit sequence learning: accounting for sequence structure and explicit knowledge", *Psychological Research*, Vol. 60 No. 1, pp. 24-41, doi: [10.1007/BF00419678](https://doi.org/10.1007/BF00419678).
- Davenport, T. and Prusak, L. (1998), "Working knowledge: how organizations manage what they know", *Ubiquity*, Vol. 1, doi: [10.1145/348772.348775](https://doi.org/10.1145/348772.348775).
- Dzedzickis, A., Kaklauskas, A. and Bucinskas, V. (2020), "Human emotion recognition: review of sensors and methods", *Sensors (Basel, Switzerland)*, Vol. 20 No. 3, doi: [10.3390/s20030592](https://doi.org/10.3390/s20030592).
- Guitart, I., Moré, J., Duran, J., Conesa, J., Baneres, D. and Gañan, D. (2015), "A semi-automatic system to detect relevant learning content for each subject", *Paper Presented at 2015 International Conference on Intelligent Networking and Collaborative Systems*, doi: [10.1109/INCoS.2015.62](https://doi.org/10.1109/INCoS.2015.62).
- Guneysu Ozgur, A., Wessel, M.J., Olsen, J.K., Johal, W., Ozgur, A., Hummel, F.C. and Dillenbourg, P. (2020), "Gamified motor training with tangible robots in older adults: a feasibility study and comparison with the young", *Frontiers in Aging Neuroscience*, Vol. 12, p. 59, doi: [10.3389/fnagi.2020.00059](https://doi.org/10.3389/fnagi.2020.00059).
- Hadavand, A., Muschelli, J. and Leek, J. (2019), "Analysis of student behavior using the R package *crsra*", *Journal of Learning Analytics*, Vol. 6 No. 2, pp. 140-152, doi: [10.18608/jla.2019.62.10](https://doi.org/10.18608/jla.2019.62.10).
- Heckman, J.J. (2008), "Schools, skills, and synapses", *Economic Inquiry*, Vol. 46 No. 3, p. 289, doi: [10.1111/j.1465-7295.2008.00163.x](https://doi.org/10.1111/j.1465-7295.2008.00163.x).

-
- Khosravi, H., Shabaninejad, S., Bakharia, A., Sadiq, S., Indulska, M. and Gašević, D. (2021), "Intelligent learning analytics dashboards: automated drill-down recommendations to support teacher data exploration", *Journal of Learning Analytics*, Early Access Articles, pp. 1-22, doi: [10.18608/jla.2021.7279](https://doi.org/10.18608/jla.2021.7279).
- Kirkpatrick, D.L. (1959a), "Techniques for evaluating training programs", *Journal of the American Society of Training Directors*, Vol. 13 No. 11, pp. 3-9.
- Kirkpatrick, D.L. (1959b), "Techniques for evaluating training programs: part 2—learning", *Journal of the American Society of Training Directors*, Vol. 13 No. 12, pp. 21-26.
- Kirkpatrick, D.L. (1960a), "Techniques for evaluating training programs: part 3—behavior", *Journal of the American Society of Training Directors*, Vol. 14 No. 1, pp. 13-18.
- Kirkpatrick, D.L. (1960b), "Techniques for evaluating training programs: part 4—results", *Journal of the American Society of Training Directors*, Vol. 14 No. 2, pp. 28-32.
- Kirkpatrick, D.L. and Kirkpatrick, J.D. (2006), *Evaluating Training Programs – The Four Levels*, 3rd ed., ISBN 978-1-57675-348-4.
- Kirkpatrick, J.D., Kirkpatrick, W.K., Kirkpatrick, D.L. and Biech, E. (2016), *Kirkpatrick's Four Levels of Training Evaluation*, ATD Press, Alexandria, VA.
- Kuruba, M. (Ed.) (2019), *Role Competency Matrix*, Springer Singapore, Singapore, doi: [10.1007/978-981-13-7972-7](https://doi.org/10.1007/978-981-13-7972-7).
- Kyllonen, P.C., Lipnevich, A.A., Burrus, J. and Roberts, R.D. (2014), "Personality, motivation, and college readiness: a prospectus for assessment and development", *ETS Research Report Series*, Vol. 2014 No. 1, pp. 1-48, doi: [10.1002/ets2.12004](https://doi.org/10.1002/ets2.12004).
- Lalor, J., Lorenzi, F. and Rami, J. (2015), "Developing professional competence through assessment: constructivist and reflective practice in teacher-training", *Eurasian Journal of Educational Research*, Vol. 15 No. 58, doi: [10.14689/ejer.2015.58.6](https://doi.org/10.14689/ejer.2015.58.6).
- Leitner, P., Khalil, M. and Ebner, M. (2017), "Learning analytics in higher education—a literature review", in Peña-Ayala, A. (Ed.), *Learning Analytics: Fundamentals, Applications, and Trends, Studies in Systems, Decision and Control*, Springer International Publishing, Cham, Vol. 94, pp. 1-23, doi: [10.1007/978-3-319-52977-6_1](https://doi.org/10.1007/978-3-319-52977-6_1).
- Levin, H.M. (2012), "More than just test scores", *PROSPECTS*, Vol. 42 No. 3, pp. 269-284, doi: [10.1007/s11125-012-9240-z](https://doi.org/10.1007/s11125-012-9240-z).
- Loh, C.S. (2013), "Improving the impact and return of investment of game-based learning", *International Journal of Virtual and Personal Learning Environments*, Vol. 4 No. 1, pp. 1-15, doi: [10.4018/jvple.2013010101](https://doi.org/10.4018/jvple.2013010101).
- Makransky, G., Bonde, M.T., Wulff, J.S.G., Wandall, J., Hood, M., Creed, P.A., Bache, I., Silahatoglu, A. and Nørremølle, A. (2016a), "Simulation based virtual learning environment in medical genetics counseling: an example of bridging the gap between theory and practice in medical education", *BMC Medical Education*, Vol. 16, p. 98, doi: [10.1186/s12909-016-0620-6](https://doi.org/10.1186/s12909-016-0620-6).
- Makransky, G., Thisgaard, M.W. and Gadegaard, H. (2016b), "Virtual simulations as preparation for lab exercises: assessing learning of key laboratory skills in microbiology and improvement of essential non-cognitive skills", *PloS One*, Vol. 11 No. 6, e0155895, doi: [10.1371/journal.pone.0155895](https://doi.org/10.1371/journal.pone.0155895).
- Mayrhofer, W., Nixdorf, S., Fischer, C., Zigart, T., Schmidbauer, C. and Schlund, S. (2021), "Learning nuggets for cobot education: a conceptual framework, implementation, and evaluation of adaptive learning content", *Proceedings of the Conference on Learning Factories (CLF) 2021, SSRN Electronic Journal*, doi: [10.2139/ssrn.3868713](https://doi.org/10.2139/ssrn.3868713).
- Mayrhofer, W., Schneikart, G., Fischer, C., Papa, M. and Schlund, S. (2022), "Measuring learning efficacy of training modules for cobots", *12th Conference on Learning Factories (CLF)*, Singapore, April, pp. 11-13.
- Roe, R.A. (2002), "What makes a competent psychologist?", *European Psychologist*, Vol. 7 No. 3, pp. 192-202, doi: [10.1027/1016-9040.7.3.192](https://doi.org/10.1027/1016-9040.7.3.192).

- Schneikart, G. (2020), "Recent trends in the development of assistance systems for biomedical research from a managerial perspective", doi: [10.34726/hss.2020.80760](https://doi.org/10.34726/hss.2020.80760).
- Schon, D.A. (1983), *The Reflective Practitioner How Professionals Think in Action*, Temple Smith, London.
- Schon, D.A. and DeSanctis, V. (1986), "The reflective practitioner: how professionals think in action", *The Journal of Continuing Higher Education*, Vol. 34 No. 3, pp. 29-30, doi: [10.1080/07377366.1986.10401080](https://doi.org/10.1080/07377366.1986.10401080).
- Siemens, G. (2013), "Learning analytics: the emergence of a discipline", *American Behavioral Scientist*, Vol. 57 No. 10, pp. 1380-1400, doi: [10.1177/0002764213498851](https://doi.org/10.1177/0002764213498851).
- Wong, B.T.-M., Li, K.C. and Choi, S.P.-M. (2018), "Trends in learning analytics practices: a review of higher education institutions", *Interactive Technology and Smart Education*, Vol. 15 No. 2, pp. 132-154, doi: [10.1108/ITSE-12-2017-0065](https://doi.org/10.1108/ITSE-12-2017-0065).

Corresponding author

Gerald Schneikart can be contacted at: gerald.schneikart@fh-wien.ac.at