Journal of Scientific & Industrial Research
Vol. 80, December 2021, pp. 1096-1106

Performance Prediction Reliability of Computer-aided Work Simulations and Employment Tests: A Case of Selecting Blue-collar Employees for Repetitive Tasks

Tekiner Kaya

International Trade and Logistics, Faculty of Business Administration and Economics, Nevsehir Haci Bektas Veli University, 50300, Nevsehir, Turkey

Received 04 Nov 2020; Revised 29 Nov 2021; Accepted 02 December 2021

The process of selecting the right candidate may differ based on job content and process dynamics. Computer-aided work simulation assessment (CAWSA) tools and employment tests are typically used in recruitment processes to achieve good Person-Job (P-J) fit. Related to this, the paper aims to determine the effectiveness and reliability of CAWSA processes and employment tests in predicting repetitive work performance amongst blue-collar employees. Additionally, the ability of these tools to predict P-J fit for repetitive tasks is analysed. Stepwise and ordinal linear regression models were used to determine the predictive capacity of CAWSA techniques and employment tests in relation to actual repetitive work performance. The model was applied on a large-scale automotive company in Turkey. A total of 142 blue-collar candidates participated in the designed recruitment process, of whom 106 were recruited in four different shops at the factory wherein they worked on repetitive tasks for six months. The results show that 84% of the variation on actual work performance can be explained by five different types of CAWSA tools, while employment tests are unable to produce the same results. Finally, a strong correlation (71%) between six months of shop performance and related shop-specific CAWSA process performance is observed, indicating that CAWSA processes can ensure effective P-J fit for repetitive tasks.

Keywords: Assessment tools, Blue-collar recruitment, Ordinal linear regression, Repetitive work, Stepwise linear regression

Introduction

Human resource (HR) is a crucial function of organisational survival, because success of an organisation is confined within the limits of its human capital.[1] The selection of suitable employees for any organisation requires an intricate combination of systems that complement one another. First, effective systems must attract a substantial pool of suitable applicants.[2] The nature of processes also differ based on job content, product, scale and volume, performance of employees, qualification, harmonious relationship amongst employees, and work dynamics. Thus, a well-designed recruitment process that meets the specific needs of firms and shops requires establishing strong pillars for corporations. Furthermore, the recruitment process must be able to measure candidates' work performances specific to the job content.

The work performed by blue-collar employees differs considerably from that of white-collar employees. Blue collar work is often prescribed in specifications and general trade practices. Often, some experience or occupational knowledge and few work-related skills are required for blue-collar employment. Thus, procedures utilized for selection of blue-collar employees have very limited range.[3] 'Blue-collar employees' refer to those whose employment requires them to perform manual labour for a good amount of time, and when the work involves the use of machines, for example, the employee must possess a good amount of skills to operate them.[4] Generally, labour-intensive industries employ huge numbers of blue-collar employees.

The ability to select and hire talented blue-collar employees for labour-intensive organisations is crucial to the success of a firm. To ensure recruitment quality, basic dimensions should be considered, including corporate image, human resources systems, management approach and company culture, in addition to recruitment activities and tools. The Assessment Centre (AC) methodology is a very popular approach in development, hiring and promotion processes. However, ACs are generally structured and used for selecting white-collar workers and managerial positions.[3] It is generally assumed to possess high levels of content-related validity and

Author for Correspondence
E-mail: tekiner.kaya@nevsehir.edu.tr

result in favourable applicant reactions compared to cognitive ability tests.[5] Two of the main assessment tools used in blue-collar recruitments are paper-online-based multiple-choice exams (POBME) and on-the-job assessment models. However, predicting the work performance potential of computer-aided work simulation assessment (CAWSA) systems and POBME may differ based on process dynamics. Hiring blue-collar employees who perform well at POBME (e.g., ability or aptitude tests) may cause motivational problems and falling job satisfaction rates later on, especially if they are required to work on routine repetitive tasks for several years (even at the beginning of their career). Thus, for a company to select appropriate employees for repetitive tasks, it should focus on repetitive task performance predictors during initial assessment.

Repetitive or routine tasks, such as assembling, connecting wires, painting and welding, along with vehicle assembly processes, have a very different nature compared to other tasks as they are usually considered stressful and monotonous and may potentially harm the employees. The Occupational Safety and Health Administration (OSHA) describes 'repetition' as 'performing the same motion or a series of comparable motions frequently or for an extended period of time.[6] A 'repetitive task' is defined as performing the same motion or series of motions continually or frequently for an extended period of time.[7] The Health and Safety Executive (HSE) report (HSE, 2010) defines 'repetitive work' as consisting of a sequence of upper limb actions of fairly short duration, which are repeated over and over again and are almost always the same (e.g. manufacturing an item, packaging a part or stitching a specific cloth).[8] Companies focus on selecting the most talented ones while considering the organisational needs for the future. As a result, the selection process should be able to measure several types of ability and competence such as the ability to work on repetitive tasks for years, leadership, resilience, and perseverance.

Considering the huge number of repetitive processes in the manufacturing plants that operate in labour intensive sectors such as the automotive, petro chemistry, metal and textile industries, and recruitment tools should be able to select appropriate candidates who are good at repetitive works. Specifically, takt time-based pull production control systems need employees who have ability to work on repetitive, tedious and patience-requiring processes. Instead of selecting the most talented employees for that type of work environment, selecting the most appropriate candidates will be able to generate more productive and efficient results. Although typical assessment exercises utilise a variety of measurement methods to tap into all facets of competency[9], work performance predictions completed by calculating an overall assessment may not be an effective way to make a hiring decision. Thornton & Rupp (2012) explained this situation as some of the dimensions being expected to be more cognitively or non-cognitively loaded than others.[10] Arthur (2012) expressed that given the centrality of the AC scores that summarise dimensional performance across all exercises in relation to AC practice, the limited attention of this unit to scoring is a notable deficiency in the literature.[11] Hoffman *et al.* (2007) also emphasises the differences between organisational behaviour and the technically-oriented aspects of task performance in terms of the job performance aspect.[12]

Aptitude or ability tests are used to look at a candidate's behaviour, intelligence and cognitive nature. While aptitude tests try to predict learning and the general ability to process information, ability tests tend to measure job-related intellectual characteristics. For instance, verbal reasoning, meaning, simple checking, and complex reasoning tests are used to select the best employees in line with the company priorities. A high number of commercially used tests have been questioned in terms of their fitness for use in time but without any clear responsibility, hence there is little chance of remedial action.[13] Feltham *et al.* (1994) state that wrong selection decisions should be more commonly attributed to an inappropriate use of tests than to the tests themselves being inherently bad.[14]

On the other hand, Thornton & Rupp (2006) claim that interviews and POBME tests, which were commonly a part of early ACs, can still provide valuable information.[15] The performance prediction power of these tests might change based on the test design, on their application or on the type of tasks involved. Finding out the prediction power of POBME and on the job assessment activities for repetitive works will be beneficial for decision-makers. Ryan *et al.* (2015) states that employment tests are typically used along with other tools to make selection decisions.[16] They also exhibit that personality (84.5%), abilities (81.6%), and leadership competencies (65.3%) are the most common characteristics assessed by tests. Employment tests such as math, verbal, and language are commonly used to assess abilities and they seem to be an efficient way of predicting work performance. However, the validity of these tools to measure repetitive or routine tasks should be checked.

An integrated AC methodology uses different tools in combination with each other and can involve different types of assessment tool such as leaderless group discussions, role play, in-basket, case analysis, structured interviews, personality and cognitive ability tests and in-depth simulation exercises. The overall assessment rating is an evaluation of the overall assessment performance. Thornton & Rupp (2006) exhibits that on-the-job evaluation exercises that should simulate organisational situations are one of the distinguishing tools that have ability to predict work performance well.[15] A detailed replica of a select task may be a good subject in this type of exercise. Since high technology usage in recruitment processes and computer aided simulation process increase the realism of the exercises and assure the recruitment of qualified blue-collar members, companies have started to invest in this area. Different recruitment exercises, such as on the job assessments or simulation/role-play exercises in which applicants come up with ideas to solve problems, have a strong correlation with successful job performance[17] and high validity.[18] CAWSA processes have also been structured to measure soft and hard skills for operational, managerial, promotional and developmental purposes. Good CAWSA practices provide an objective selection system, a good preview of the job level, and good work performance prediction benefits.

While assessment activities focus on determining the most appropriate candidates for corporations, they also have the capability to assign candidates to an appropriate process which is known as the person-job fit (P-J). Carless (2005) also emphasises that applicants who perceive a fit between their knowledge, skills and abilities and the job requirements are likely to remain engaged in the selection process and accept a job offer.[19] In addition, good established P-J fit will result in high employee engagement,[20] job satisfaction and reduced turnover intentions.[21]

In this study, predicting the work performance reliability of the different recruitment tools for repetitive and specialised processes in labour intensive sectors for blue-collar employees has been analysed. These three tools are the general POBME, competency-based designed POBME (in six different dimensions), and CAWSA. Additionally, by assigning the right candidate to the right shop (known as P-J fit), the effectiveness of CAWSA is exposed via an assessor's evaluation and four different CAWSA scores.

This paper structured as follows: the materials and methods section introduce the research design and methodology used to figure out the performance of the assessment tools and processes for repetitive works; the findings and results are presented in the results and discussion section, followed by the limitations and conclusions.

Materials and Methods

Research Design

The advancement of digital technology has forced many human resource processes to adopt knowledge-based technological tools to support their daily operations such as software for selecting talented people, assessing the workforce performance, and promoting talented employees. Since recruitment is one of the most significant functions of HR, it plays a critical role in the sustainable performance of organizations. Many companies are bringing computerised systems and products into their HR recruitment system. To measure the work performance prediction power of the three specified assessment tools, a new competency-based work specific POBME and CAWSA process has been designed at the Toyota Motor Turkey (TMT) plant. TMT uses three different packaged paper-based employment tests for maths, meaning and reading with 20 questions each. The tests are focused on measuring technical abilities.

In line with the recruitment improvement studies conducted at TMT, the company developed 6 different TMT-specific competency-based POBMEs with a consultancy company with experience of employment test design. The aim of the tests was to expose the candidates' fundamental skill potential based on the TMT competencies. Twenty custom-made questions were designed for each test including meaning, analytical thinking, analytical comparison, virtualisation, problem solving and virtual mind. TMT additionally configured four shop-specific CAWSA processes; welding, paint, logistics and assembly. These processes are the key processes that make up 85% of the total hiring. Each CAWSA process, specifically spot welding (for welding), weight mounting (for logistics), sealer application (for paint) and bolt tightening (for assembly) has been established based on the competencies required by each shop. The CAWSA processes measure the performance of the candidates via checking the number of correct cycles, the number of total cycles, and the total duration. The system controls the whole process via components such as the laser sensors for the welding operations positioning, on/off switches for bolt tightening and the logistics operations, connectors and circuit checkers for welding and sealer operations, start/stop, emergency buttons, and Andon. Candidates wear safety shoes, a helmet, safety jacket, earflaps, and safety gloves as usual in the manufacturing area and they work for 1 hour per process. In total, they spend four hours engaged in the CAWSA process where the manufacturing area noise is also simulated. In Fig. 1, how the CAWSA processes were established to measure and predict repetitive work performance in TMT is presented. To assist the assessor and to give the necessary orders, LCD screens were used. PLC Connection Box coordinated all of the sensors and PCI I/O Cards performed the sensor communication. An SQL Server was used to collect and store the data. Visual Studio VB.NET was selected as the software development environment because of the PLC software Dll's compatibility with Visual Basic. Two main programs were used in the CAWSA processes. The first software for administration purposes was used to control, manage, monitor, and track the real time assessments and reporting. Second, there was the software for the four different CAWSA processes.

TMT employs more than 4000 blue-collar employees and each month they hire more than a hundred temporarily contracted (six months) employees. The primary objective of the recruitment process is to predict the suitability for repetitive work performance of the blue-collar workers rather than their leadership, self-reliance and decisiveness competencies. The reason for this is that the HR grade system makes it mandatory for new blue-collar employees to work in operations (generally engaged in repetitive processes) for at least 10 years before promoting them to management level as either a team leader or group leader. The company assesses the

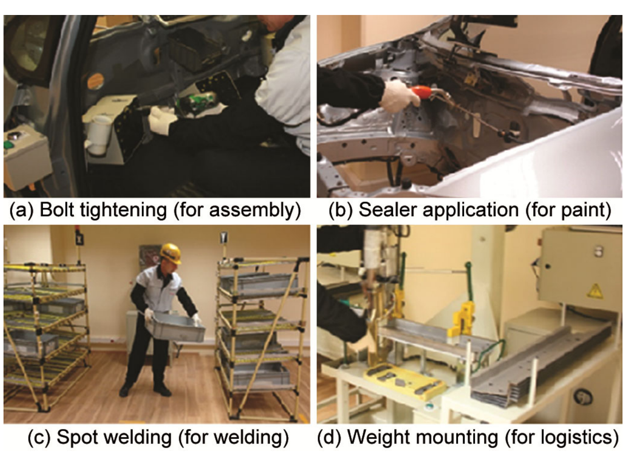

Fig. 1 — CAWSA processes to assess repetitive work performance

employees during this period and promotes the most talented ones.

The research was conducted in three sections as shown in Fig. 2: the recruitment processes, the hiring process, and a work performance evaluation after six months of working. First, the candidates attend the TET tests. Without checking their performance at TET, they proceed to the interview. If the interviewer observes a serious problem that will affect either the safety or security standards, the candidates fail. Otherwise, they continue with the recruitment process. In steps 3 and 4, the candidates attend the new competency-based POBME and CAWSA. If the CAWSA assessor observes a serious problem in connection to the candidate or if the candidates give up during the CAWSA process, they fail. After CAWSA, the candidates begin the health check process. If they pass, they are recruited. They are assigned to four different shops at TMT and work for six months. The assignments done by the assessor are based on his observations and the candidate's CAWSA process performance. After six months of working at TMT, their work performance is evaluated by their leader (team or group leader). This process was designed to measure repetitive work performance and the prediction power of the three different assessment tools.

Data

As shown in Fig. 2, a total of 142 candidates were selected to attend the process in line with the research design at the TMT plant. Twenty-eight candidates left during CAWSA and four candidates were eliminated by the assessor due to misbehaviour and safety problems. During the statistical analysis, four outliers were observed and removed. Overall, 106 candidates were recruited. The descriptive statistics about the candidates' performance during the recruitment processes is presented in Table 1. Work performance expresses the real work performance scores of the employees in the manufacturing area after six months of working. The TET and competency-based POBME scores are presented out of 100.

Data Analysis Methodology

The next step in the research consisted of determining the assessment efficiency of the three main recruitment tools for repetitive works, specifically an evaluation of the following multiple model approach. The stepwise linear regression (SLR) model, which is a combination of the forward and backward selection

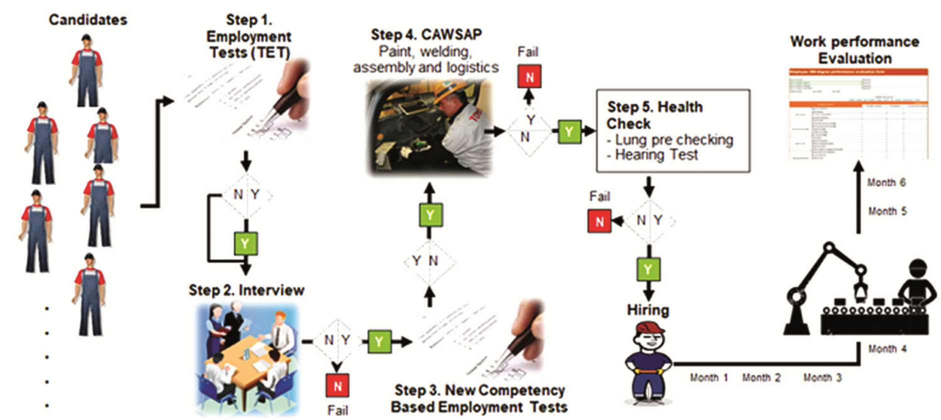

Fig. 2 — Research design

Table 1 — Descriptive statistics of candidates' performance on actual work, POBME and CAWSA

Recruitment Tools	Performance on	Mean	Std Dev
On the Job	Work Performance	85.87	7.07
Ex-Employment Tests	TET (Math, meaning and reading)	46.92	13
Competency Based POBME	Analytical Comparison	52.37	15.98
	Meaning	49.09	18.82
	Virtual Mind	51.58	19.34
	Virtualization	35.91	26.02
	Analytical Thinking	49.68	14.98
	Problem Solving	45.11	20.62
New CAWSA Processes	Bolt tightening number of cycle	182	24
	Bolt tightening percentage of successful cycle (%) (%)	86	6,7
	Sealer application number of failure per cycle	17	6
	Welding number of cycle	38	6
	Welding percentage of successful cycle (%)	73	6
	Weight mounting number of cycle	242	41
	Weight mounting percentage of successful cycle (%) (%)	88	6

techniques, was used to determine the prediction power of the recruitment tools for repetitive works. Actual work performance after six months of working through the TMT processes were selected as the dependent variable while the 14 recruitment tools seen in Table 1 were selected as the independent variables.

It is assumed that there is a linear relationship between the dependent variable and independent variable in the linear regression (LR) model. For the data of the n statistical units set, $\{y_i, x_{i1}, \ldots, x_{ip}\}_{i-1}{}^{n}$, the relationship are modelled via error variable ε represented as:

$$y = X\beta + \varepsilon \qquad \ldots (1)$$

where

$y = [y_1, y_2 \ldots, y_n]T, X = [x_1T, x_2T \ldots, x_nT]T, \beta = [\beta_0, \beta_1 \ldots, \beta_p]T \ and \ \varepsilon = [\varepsilon_1, \varepsilon_2 \ldots, \varepsilon_n]\,T.$[22]

The number of independent variables is equal to the length of β in the LR model. The LR focuses on obtaining β. Many estimation methods with different computing complexities were used to expose β. One of the LR methods in which step-by-step iterative construction utilised is SLR. SLR selects the independent variables automatically and is constructed by adding or deleting independent variables via the forward selection or backward elimination procedures.[22]

Since the aim of this study was to determine performance predictors for repetitive works, all of the candidates' variables in the model were checked to see the significance level via comparing it with specified tolerance level. In case a non-significant variable was observed, it was removed. However, because of the improper research goals or due to a lack of information, there are many instances where a statistical power analysis is not reliable. Green (1991) states that $N \geq 50 + 8$ m for a multiple correlation test. m represents the number of variables while N is the sample size.[23] Khamis & Kepler (2010) used this as the criterion for developing a formula for the minimum sample size in a multiple regression model with continuous predictors.[24] The ultimate formula, $N = 20 + 5k$ where k represents the number of predictors, is simple and optimal with respect of the principle based on the rate of change of the reliability criterion relative to n. In this study, the minimum sample size should be 90 (n = 20 + 5.14). This means that the model is reliable and is statistically powerful.

Four general steps are used to set the SLR model to help determine the significant predictors. The four assumptions are normality, linearity, heteroscedasticity and multicollinearity. These were checked and no unconformity was observed.

After determining the significant repetitive work performance predictor recruitment tools via SLR, the ordinal logistics regression (OLR) model was utilized to determine the strength of the predictors which were categorized into three groups: very good, average, and below average. The OLR model is selected since OLR has the ability to expose how the significant predictors influence repetitive work performance. When the distance between the different categories of response has a clear sense of ordering, the response is considered to be ordinal. The model is as follows for an ordinal response:

$$\hat{A} = x_i\beta + \varepsilon_i \quad \dots (2)$$

In Eq. (1), $\hat{A}$ is the linear function that determines the discrete outcome while x_i represents a vector of the observable features, and β indicates a vector of the regression coefficients. Finally, ε_i demonstrates an error term with a logistic distribution with a mean of zero and a variance of $\pi^2/3$.

One of the assumptions of OLR is the proportional odds. This means that each independent variable has an identical effect at the point of each cumulative split of the ordinal dependent variable.

In addition to figuring out the real repetitive work performance recruitment indicators, TMT requested clarification on whether the significant work performance predictors differ based on the graduation school type of the candidates. The school types tested in this study were high school and occupational high school. The independent sample - t test was applied to check the hypothesis. The independent sample -t test assumptions which are the normally distributed data, the lack of significant outliers, the homogeneity of variance and the continuous scale measurements were checked.

The final research question asked how efficient the statistically significant repetitive work performance predictors are for determining P-J fit. The capability of the CAWSA processes when it comes to assigning the right candidates to the right processes was analysed. After step 5 in Fig. 2, the assessor decides where the successful candidates will work (in which shop) based on his observations and the CAWSA results. The assessors additionally consider the process-based CAWSA scores of the candidates in addition to the overall assessment performance. For instance, the candidates who performed best in the bold tightening process in CAWSA were assigned to the assembly shop. The Pearson correlations between the relevant shop work performance and the score of the CAWSA process that appoints the candidates to a particular shop were checked to measure the effectiveness of the P-J fit. A high correlation was interpreted as an effective P-J fit. The multiple model approach is presented in Fig. 3.

Results and Discussion

Reliability

The assumptions for SLR, the OLR model and the independent-t test were tested. There are many types of normality test that have been used in the literature such as D'Agostino's statistics, the Shapiro-Wilk test, and Anderson-Darling. Zhang & Wu (2005) admit that the Shapiro-Wilk test, which is essentially the squared ratio of the best linear unbiased estimator for scale to the standard deviation, is one of the most powerful tests of normality.[25] The test results show that the data is normally distributed ($p > 0.05$). Multicollinearity occurs when the independent variables in a regression model are correlated. This is a problem since the regression analysis must isolate the relationships between each of the dependent variables and the independent variable. In order to detect whether there were any similarities between the independent variables, the coefficient correlations were checked (Table 2). Since strong relationships were observed ($r > 0.80$), the collinearity test results are stable and no multicollinearity was observed (the VIF values vary from 1.05 to 2.4).

The linearity and heteroscedasticity assumptions were checked via scatterplots. The SLR model has a strong prediction power for repetitive work performance

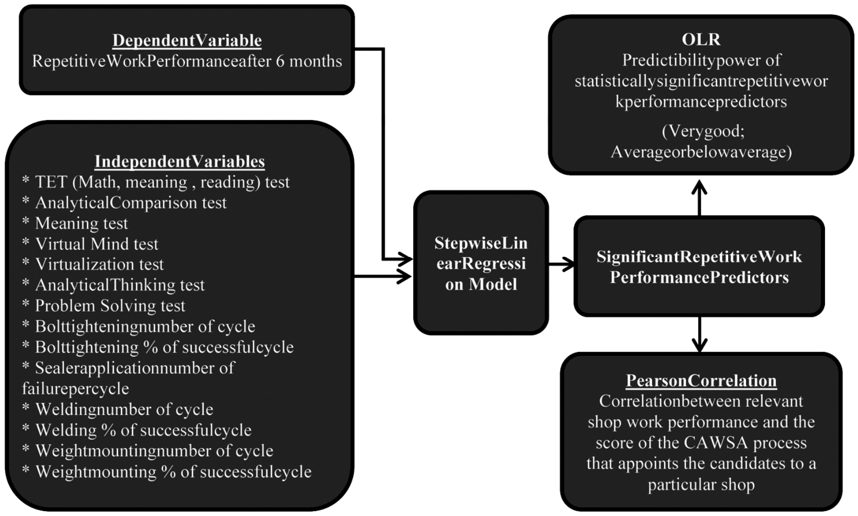

Fig. 3 — Multiple research model design

Table 2 — Coefficient correlations of assessment tools (%)

Recruitment Tools	Job Performance	TET	Analytical Comp	Meaning	Virtual Mind	Visualization	Analytical Thinking	Problem Solving	Bolt Tightening # of Cycle	Bolt Tightening Succ. Perc. (%)	Sealer Faulty	Weld # of Cycle	Weld Succ. Perc (%)	Weight Mounting # of Cycle	Weight Mounting Succ. Perc. (%)
Job Performance	100	−2	−11	−09	1	−3	1	1	81	61	−68	77	70	78	55
TET		100	58	62	72	78	59	60	05	4	8	05	6	−4	10
Analytical Comp			100	30	32	34	24	23	−13	−7	13	04	−11	−16	9
Meaning				100	34	24	34	34	3	−4	18	−03	−6	−7	11
Virtual Mind					100	46	46	28	5	9	2	12	19	11	24
Visualization						100	39	31	8	9	0	08	9	−5	1
Analytical Thinking							100	8	3	7	17	−8	4	−8	11
Problem Solving								100	5	−3	−6	−2	0	4	−8
Bolt Tightening # of Cycle									100	49	−54	67	58	65	44
Bolt Tightening Succ. Perc (%)										100	−41	56	46	50	28
Sealer Faulty											100	−66	−65	−67	−31
Weld # of Cycle												100	55	68	56
Weld Succ. Perc. (%)													100	61	36
Weight Mounting of Cycle														100	48
Weight Mounting Succ. Perc (%)															100

in an accurate manner. A rectangular pattern in the dots was observed when the relationship between the standardised residuals and the standardised predicted value was checked (Table 3). The SLR model results are considered to be significant and were accepted. The R square change also shows that 83.9% of the variance in the work performance can be explained by the six statistically significant predictors. This shows that a fairly good model fit has been achieved.

By using parallel lines, the assumption of proportional odds was tested for the OLR model. Since the significance of the Chi-Square statistics was observed 0.838 which is greater than 0.05, the proportional odds assumption appears to have held. The

Table 3 — Model summary of stepwise linear regression

Model	R	R^2	Adjusted R Square	Std. Error of the Estimate	Change Statistics		
					R^2	F Change	Sig. F Change
1	0.81(a)	0.65	0.65	4.18	0.654	181.13	0.000
2	0.88(b)	0.77	0.76	3.46	0.112	45.57	0.000
3	0.89(c)	0.80	0.79	3.20	0.035	16.60	0.000
4	0.91(d)	0.82	0.81	3.05	0.021	10.93	0.001
5	0.91(e)	0.83	0.82	2.98	0.009	5.09	0.026

Predictors: Bolt Tightening # of cycle, Weight Mounting # of cycle, Welding # of cycle, Welding successful cycle (%), Bolt Tightening successful cycle (%)

Table 4 — Coefficients of assessment tools

	Unstnd Coeff.		Stnd Coeff.	Sig.	Correlations		Collinearity Statistics
	B	Std. Error	Beta	Part	Partial	Part	VIF
(Constant)	22.23	4.56		0.000			
Bolt Tightening # of cycle	0.103	0.018	0.356	0.000	0.512	0.239	2.222
Weight Mounting # of cycle	0.037	0.011	0.217	0.001	0.326	0.139	2.460
Welding # of cycle	0.254	0.075	0.222	0.001	0.336	0.143	2.414
Welding successful cycle (%)	0.212	0.065	0.185	0.002	0.323	0.137	1.821
Bolt Tightening successful cycle (%)	0.132	0.055	0.127	0.018	0.245	0.101	1.572

goodness of model fit results, which are useful for determining whether a model exhibits a good fit with the data, present that the deviance test results were non-significant (x^2 = 72.13, p = 1.00). This means a good model fit. The continuous scale assumptions, homogeneity of variance and normality were checked using regression analysis and an independent t-test. Four significant outliers were detected and removed.

Predictive Capacity of Repetitive Work Performance

The SLR model results show that five out of the 14 recruitment tools have a statistically significant effect on repetitive work performance. These are bolt tightening number of cycles, weight mounting number of cycles, weld number of cycles, weld percentage of successful cycles (%), and bolt tightening percentage of successful cycles (%) (Table 4). This result accounts for about 84% of the variation which is also statistically significant. This shows that CAWSA-based assessments have a big role to play in predicting repetitive performance. R square changes and β coefficient indicate that the bold tightening number of process is the strongest predictor for repetitive works.

The SLR model in Eq. (3) details the final equation for predicting repetitive work performance:

$$Y(Work\ Performance) = 22.23 + 0.103\ (Bolt\ Tightening\ \#) + 0.037\ (Weight\ Mounting\ \#) + 0.254\ (Weld\ \#) + 0.212 (Weld\ Succ.Perc) + 0.132\ (Bold\ Teightening\ Succ.Perc) \quad (3)$$

In case a candidate achieves 10 more cycles in the weight mounting process, the actual work performance is going to improve by 0.37 out of 100. Similarly, 10 more bold tightening number of cycles will result in 1.03, %10 more welding successful cycle percentage 2.1 and 10 more welding number of cycle will result in 2.5.

The strength of the work performance predictors is categorized into three groups: below average, average and very good. To expose the best predictor/s, the SLR model which was used to determine statistically significant predictors is utilized in the OLR model. The pseudo R-square values are treated as rough analogues to the R^2 value in OLR. In literature, there is no common strong guidance on how these should be interpreted.[26,27] However, the pseudo R-square values do not have same interpretation as R^2. For each of the independent variables, significance test results and OLR coefficients are exhibited in Table 4. The coefficients are interpreted as the predicted change in the log odds as a higher per unit increase in the independent variable. The statistically significant dependent variables are the weight mounting number of cycles, the bolt tightening number of cycles and the welding percentage of successful cycles (sig. <0.05). This means that as the scores go up for an independent variable, there is an increased probability of the dependent variable falling at a higher level. For instance, where there is a one unit increase in the bolt tightening number of cycles, there is

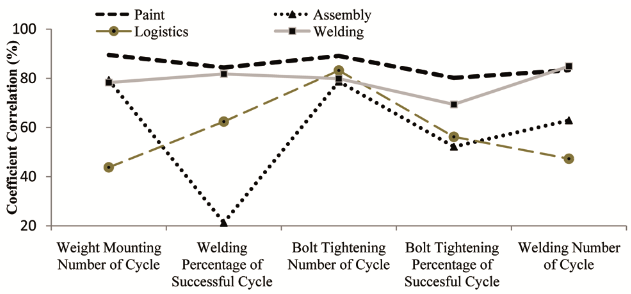

Fig. 4 — Correlations between shops and significant predictors

a predicted increase of 0.088 in the long odds of work performance at a higher level. As a result, three out of the five predictors have a significant role in differentiating between the different repetitive work performance levels.

To reveal the effect of graduated school type on predictors of repetitive work performance, the independent sample - t test was utilized. 54 of 106 candidates were graduated from occupational high school while others were from high school. The results show that the occupational high school graduates have a better repetitive work performance than the high school graduates (the sig. value is 0.038).

In order to reveal the efficiency of the significant recruitment tools in relation to how P-J fit works, the Pearson correlations between significant repetitive work performance predictors and related shop performance were checked. The coefficient correlation results in Fig. 4 show that the bolt tightening number of cycle performance is an efficient indicator of the P-J fit for all shops. This means that candidates who have a good result for the bold tightening number of cycle indicator can work on repetitive tasks in all shops. The welding CAWSA process is not efficient indicator for P-J fit for the assembly and logistics operations as expected. On the other hand, the welding process results in CAWSA are efficient indicators of P-J fit for the welding shop as expected. The weight mounting number of cycles indicator is a bit far away to be able to predict the logistics shop performance. One of the unexpected results was that the weakest coefficient correlation of the study was observed between the logistics shop performance of the employees and the weight mounting number of cycles performance in CAWSA. The bolt tightening percentage of successful cycles indicator is a good predictor for paint and welding shop performance but not a strong predictor for assembly, which is contrary to expectations.

Limitations

The candidates performed 180 cycles on average continuously in a repetitive flow and with a high degree of caution for one hour for each CAWSA processes. This required the resources for building and administering. In terms of POBMEs, the tests have the ability to measure basic competencies such as analysing the problem in detail, understanding the relationships, and using the visual mind etc. This may work to gain an understanding of the overall performance of the candidates and to distinguish between the candidates in general. However, it is difficult to consider it as a strong tool to measure repetitive work performance even if it is designed based on the company competencies or skills. The dependent variable used in this study refers to the employees' work performance when engaged in repetitive tasks for six months. As a result, the findings should not be interpreted in the manner where the competency-based designed POBMEs have no role in predicting the blue collar employees' long-term work performance.

Conclusions

In terms of the theoretical contributions, first, this study indicates that CAWSA exercises have power when it comes to predicting the employee's repetitive work performance in blue collar tasks in the event of the well-established integration of the assessment tools and company competencies. Additionally, 84% of the variation in terms of actual work performance can be explained by the CAWSA processes. However, the POBMEs cannot estimate repetitive work performance alone even if they were designed while specifically considering the company competencies. The reason might be that POBMEs are not a part of measuring repetitive work performance for long hours (especially in terms of needing to be valid for continuous work flow in the automotive industry and mass production systems).

The second contribution of this study is the observation of a strong relationship between the shop and shop-specific-designed CAWSA process performance for repetitive tasks in terms of P-J fit. This is what the POBMEs could not do. The POBME test results appear to be a weak predictor of both the self and supervisor ratings in terms of the work performance of the blue-collar workers.

To increase the efficiency of assigning appropriate candidates to an appropriate process, considering the relative process dynamics is essential. A well-established P-J fit has the potential to result in lower turnover intentions. The findings do not indicate anything about turnover since none of the candidates resigned during the six months of work. However, 28 candidates gave up during CAWSA and four candidates were eliminated by the assessor during CAWSA out of 142 candidates. This results in a 22.5% elimination ratio. These findings show that on the job assessment processes and tools have the potential to reduce the turnover risk for blue-collar employees in relation to repetitive tasks. The strong correlation (71% on average) between shop-specific CAWSA performance and the assigned shop work performance exhibits that on-the-job assessment activities have the power to ensure an effective P-J fit.

Future Directions

Further studies may focus on the role of CAWSA in terms of the turnover ratio. Besides this, the working hours for the CAWSA processes might play a critical role in the power of the work performance prediction. Because of the high cost, companies generally do not prefer long hours of assessment for blue-collar employees because of the cost. Instead of a one-hour assessment per CAWSA, various assessment hours may be applied and the total costs can be compared. Additionally, the repetitive work performance of the candidates after six months of working might be analysed in terms of skills, knowledge, motivation, situational constraints and abilities. The relations between CAWSA-POBME and their dimensions may be additionally analysed.

Acknowledgement

This work and research environment is partially supported by Toyota Turkey.The initial findings of this project are first presented in 18th and 19th International Business Congress in Turkey by Kaya (2019) and Kaya & Karacaoglu (2020).

References

1 Phillips J M & Gully S M, Multilevel and strategic recruiting: Where have we been, where can we go from here? *J Manage*, **41(5)** (2015) 1416–1445.

2 Chapman D S & Mayers D, Recruitment processes and organizational attraction, in *Employee Recruitment, Selection, and Assessment: Contemporary Issues for Theory and Practice* edited by I Nikolaou & J K Oostrom (Psychology Press, London) 2015, 27–40.

3 Carless S A, Fewings-Hall S, Hall M, Hay M, Hemsworth P H & Coleman G J, Selecting unskilled and semi-skilled blue-collar workers: The criterion-related validity of the PDI-Employment Inventory, *Int J Sel Assess*, **15** (2007) 335–340.

4 Melamed S, Ben-Avi I, Luz J & Green M, Objective and subjective work monotony: Effects on job satisfaction, psychological distress and absenteeism in blue-collar workers, *J Appl Psychol*, **80(1)** (1995) 29–42.

5 Macan T H, Avedon M J, Paese M & Smith D E, The effects of applicants' reactions to cognitive ability tests and an assessment center, *Pers Psychol*, **47(4)** (1994) 715–738.

6 OSHA (Occupational Safety and Health Administration) Publication, *Solutions for the Prevention of Musculoskeletal Injuries in Foundries*, (2012) 3465–3508. [https://www.osha.gov/Publications/osha3465.pdf. (7 August 2020)].

7 OSHA (Occupational Safety and Health Administration) Publication, Guidelines for Retail Grocery Stores: Ergonomics for the Prevention of Musculoskeletal Disorders, (2004) 3192 [https://www.osha.gov/ergonomics/guidelines/retailgrocery/retailgrocery.html. (15 August 2020)].

8 HSE (Health and Safety Executive) Report, *Assessment of repetitive tasks of the upper limbs (the ART tool): Guidance for Employers*, (2010) Leaflet Indg438, HSE Toolkit. [www.hse.gov.uk/pubns/indg438.pdf. (5 August 2020)]

9 Thornton G C III, Rupp D E & Hoffman B, *Assessment Center Perspectives for Talent Management Strategies*, (Routledge, New York) 2015, 3–24.

10 Thornton G C III & Rupp D R, Research into dimension-based assessment centers, in *The Psychology of Assessment Centers* edited by D J R Jackson, C E Lance & B J Hoffman, (Routledge, New York) 2012, 141–167.

11 Arthur W Jr, Dimension based assessment centers: Theoretical perspectives, in *The Psychology of Assessment Centers*, edited by D J Jackson, B J Hoffman & C E Lance (New York) 2012, 96–120.

12 Hoffman B J, Blair C, Meriac J & Woehr D J, Expanding the criterion domain? A meta-analysis of the OCB literature, *J Appl Psychol*, **92** (2007) 555–566.

13 Edenborough R, *Assessment Methods in Recruitment, Selection & Performance. A Manager's Guide to Psychometric Testing, Interviews and Assessment Centres*, (Kogan Page Limited, UK) 2005 27–30.

14 Feltham R, Baron H & Smith P, Developing fair tests, *Psychologist*, (1994) 23–25.

15 Thornton G C III, & Rupp D R, *Assessment Centers in Human Resource Management: Strategies for Prediction, Diagnosis, and Development*, (Lawrence Erlbaum Associates, New Jersey) 2006, 6–9.

16 Ryan A M, Inceoglu İ, Bartram D, Goluovich J, Grand J, Reeder M, Dereus E, Nikolaou I & Yao X, Trends in testing: highlights of a global survey, in *Employee Recruitment, Selection, and Assessment : Contemporary Issues for Theory and Practice*, edited by I Nikolaou & J K Oostrom, (Psychology Press, Taylor & Francis Group) 2015, 136–153.

17 Meriac J P, Hoffman B J, Woehr D J & Fleisher M, Future evidence of validity of assessment center dimensions: A meta-analysis of internal and external correlates, *J Appl Psychol*, **93** (2008) 1042–1052.

18 Dilchert S & Ones D S, Assessment centre dimensions: Individual differences correlates and meta-analytic incremental validity, *Int J Sel Assess*, **17** (2009) 254–270.

19 Carless S A, Person-job fit versus person-organization fit as predictors of organizational attraction and job acceptance intentions: A longitudinal study, *J Occup Organ Psychol*, **78(3)** (2005) 411–429.

20 Gerards R, Grip A & Baudewijns C, Do new ways of working increase work engagement? *Pers Rev*, **47(2)** (2018) 517–534.

21 Kröll C & Nüesch S, The effects of flexible work practices on employee attitudes: evidence from a large-scale panel study in Germany, *Int J Hum Resour Manag*, **30(9)** (2017) 1505–1525.

22 Peng Y, Zhang Y J & Liu D T, Degradation estimation using feature increment stepwise linear regression for PWM Inverter of Electro-Mechanical Actuator, *Microelectron Reliab*, **(88–90)** (2018) 514–518.

23 Green S B, How many subjects does it take to do a regression analysis? *Multivariate Behav Res*, **26** (1991) 499–510.

24 Khamis H J & Kepler M, Sample Size in Multiple Regression: 20 + 5k, *J App Stat*, **17(4)** (2010) 505–517.

25 Zhang J & Wu Y, Likelihood-ratio tests for normality, *Computational Statistics & Data Analysis*, **49(3)** (2005) 709–721.

26 Osborne J, *Best Practices in Logistic Regression*, (SAGE Publications, USA) 2015.

27 Pituch K A & Stevens J P, *Applied Multivariate Statistics for Social Sciences* **6th edn**, (Routledge) 2016.