International Journal of Economics, Commerce and Management
United Kingdom ISSN 2348 0386 Vol. VIII, Issue 2, February 2020
http://ijecm.co.uk/

BIASES IN SUBJECTIVE ESTIMATION OF POTENTIAL IMPROVEMENT

Pini Davidov
Department of Industrial Engineering and Management,
Azrieli - College of Engineering Jerusalem, Jerusalem, Israel
davidovpini@gmail.com

Shlomo Globerson
Faculty of Engineering, Ruppin Academic Center, Emek-Hefer, Israel,
Faculty of Management, Tel Aviv University, Israel
shlomog@ruppin.ac.il

Abstract

Executing tasks within an organization typically requires planning based on a forecast of activities and required resources. Therefore, a biased forecast is a major source for incorrect planning and faulty execution, since a biased value is consistently either above or below the real value. Identifying the bias, therefore, enables one to adjust the forecasted value and improve forecasting accuracy. This research investigates the nature of subjective bias when individuals had to forecast their potential improvement. Eighty subjects participated in the study. They were asked to repeat typing a paragraph on a computer. The actual time of the first two repetitions were given to them right after they were completed, and subjects were asked to estimate the time for the 16th repetition. The subjects underestimated their improvement potential of performance by approximately 24%. Also, performance deviation was not identical for all subjects; Performance of those with less experience was significantly higher than those more experienced. If these findings are consistent for other tasks as well, they may be used for adjusting subjective forecasts to better predict the future.

Keywords: Forecasting bias, learning curve, time estimation, subjective forecasting

INTRODUCTION

Planning is the process of identifying activities required to achieve a desired goal. The goal may be renovating a house, building a bridge, or writing a computer application. Planning involves the establishing of functional and technical specifications (customer requirements, drawings, materials etc.) of the end result, resources required (equipment and labor) for execution, sequential relationship among activities, duration required for executing each activity, and total required duration. Planning, therefore, requires forecasting of various aspects such as duration and required resources. Even if the required set of activities is carefully planned, it is important to note that very rarely can a forecast perfectly predict the future. The lower the forecasting error, the higher the chance to complete execution according to the original plan. Thus, it is of utmost importance to understand the nature of the forecasting error.

One must differentiate between "objective" and "subjective" forecasting. Objective forecasting is based on a mathematical equation in which the forecasted value (the dependent variable) is a function of other variables (the independent variables). For example, a contractor wishes to estimate labor hours required for laying a tile floor. Among other variables, labor hours are a function of variables such as total area to be tiled, tile size, and total length of the walls. It is difficult, if at all possible, to include in the regression equation all the variables that impact on the forecasted value when formulating the forecasted equation. Also, most variables are of a stochastic nature, such as time required to execute a task. Therefore, such equations will not be able to perfectly forecast the future.

Subjective forecasting is based on judgements of individuals. When performing the forecast, those individuals take into consideration their experience to give their best estimate.

When predicting task duration, Kahneman & Tversky (1979) suggest that an individual tends primarily to consider aspects of the present task they are being asked to predict, rather than past information on predictions of previous task durations and their ability to accurately predict them. This hypothesis was validated by Buehler*, et al.* (1994); when subjects were instructed to consider their previous experience, their prediction accuracy improved. Halkjelsvik and Jørgensen (2012) found that tasks take longer than predicted. In other words, individuals have a tendency to be over optimistic with regard to their future performance.

Much work has been performed on the investigation of the planning fallacy, identified by Kahneman and Tversky (1979). It states that although subjects have previous experience regarding their tendency to underestimate duration, they still continue to underestimate future performance [e.g., Buehler, R., Griffin, D., & Ross, M. (1994), Buehler, R., Griffin, D., & Peetz, J. (2010), Kruger, J., & Evans, M. (2004), Rodon, C., & Meyer, T. (2012)]. There is, however, evidence that the planning fallacy phenomenon is not consistent. For example, Thomas & Handley (2008) found that underestimation occurs with a shorter operation, and overestimation occurs with a longer one, suggesting that estimates are distorted in the direction of the anchors. König, *et al*., (2015) postulate that planning fallacy can be reduced if subjects are asked to estimate the duration of a previous task that they had performed. This improves their ability to better predict the duration of the next task, even when the nature of the future task is different from the previous one.

It is a fundamental idea in many psychological theories that a person can learn from his previous errors (i.e., via feedback). Thus, if a person makes an error and receives feedback, this decreases the probability of making the same error again. For example, Roy & Christenfeld (2008) show that informing people of how long they had previously taken on the same task, reduces prediction bias.

Despite the cited references, a literature review by Chan & Hoffmann (2017) reveals that there is limited published research concerning the estimated or subjective time required for task performance. Also, when dealing with estimation for human activities one should keep in mind the potential improvement gained with experience. Potential improvement over time as a function of repetition is depicted by a "learning curve," also known as an "improvement curve," of which the following are some of its major features.

Many learning curve models have been offered for depicting performance improvement over time [for example, Glocka, *et al*., (2018), Grossea(2015)], but the most common model is presented by equation (1)

$$(1)\ t(s) = as^{-m}$$

Where:

s – repetition number

t(s) – expected time for repetition s

a – a parameter representing the time for the first repetition

m – a parameter representing the learning progress

The following equation shows that this “power model” yields a constant percentage of reduction in performance time, for each doubling of the cumulative production:

$$(2)\ SL = \frac{\mathrm{t(2s)}}{\mathrm{t(s)}} = \frac{a(2s)^{-m}}{as^{-m}} = 2^{-m}$$

Where:

SL – the slope of the curve that expresses the constant proportion of time reduction.

For example, a slope of SL=0.8 means that for every doubling of repetition (e.g., from 4 to 8 repetitions) the time for higher repetition is only 0.8, or 80%, of the base repetitions. The learning curve slope is commonly used when companies discuss the rate of production improvement.

The objective of this study is to investigate the ability of individuals to estimate their improvement potential as expressed by their learning curves.

RESEARCH METHODOLOGY

Subjects participate in the study were 80 students between the ages 22-29, who were paid to participate in the experiment. They were requested to type the following sentence for sixteen times: “Nevertheless, something about Miss Dykstra put me on my guard.”

A computer program, developed for this study, allowed each subject to start typing the next repetition, only if the last repetition was typed correctly and without any mistakes. Otherwise the subject had to make corrections until the typing was perfect. Subjects were given the first two repetitions typing time, and they were asked to estimate the time for the 16th repetition. They then continued to repeat typing the sentence until they reached the 16th repetition. The program collected actual time for all sixteen repetitions and the time estimation for the 16th repetition, generating the database used for analysis.

A transformation of equation (1) to the logarithm domain, as presented by Equation (3), is of the following linear nature.

(3) $\log(t(s)) = \log(a) - m*\log(s)$

That is, it is possible to use a linear regression analysis on the logarithm of the data set to fit a linear line and estimate the value of the parameters, a and m, that will generate the best fit of the line to the data points.

Using regression analysis on the sixteen data points, learning curve parameters (a, m) and other related variables, reported in the analysis and results section, were established for each participant.

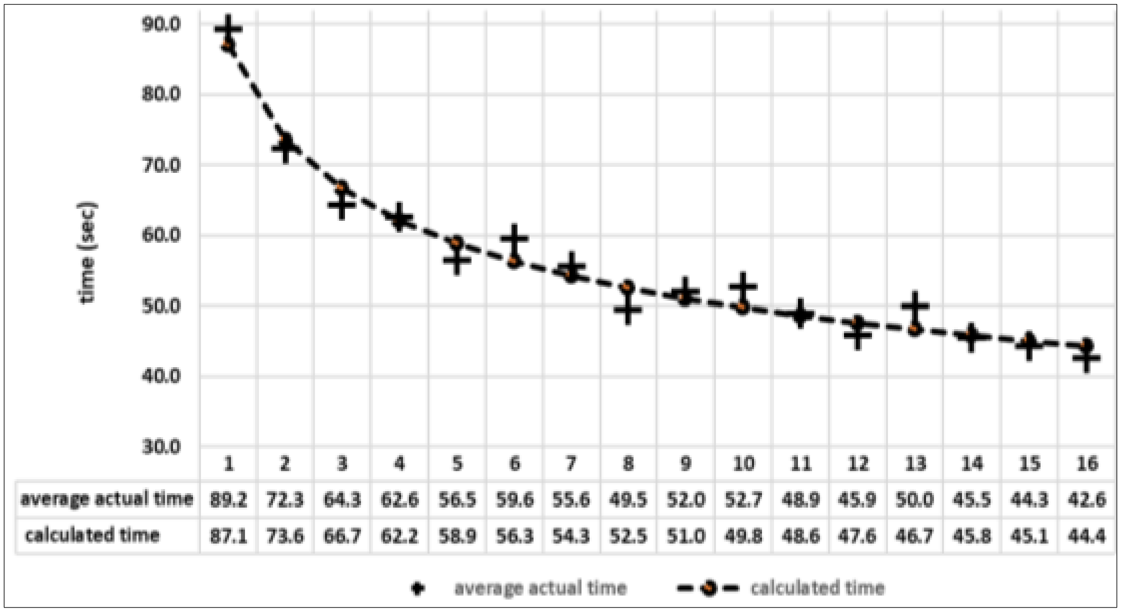
The sample used for this study consists of 80 subjects. All were undergraduate students paid to participate in the study.

ANALYSIS AND RESULTS

General

The average time to complete each of the sixteen repetitions was calculated and is presented in figure 1 to present overall behavior. Fitting a line to the 16 averages via regression analysis resulted in a learning curve, as presented by the dashed line in the figure.

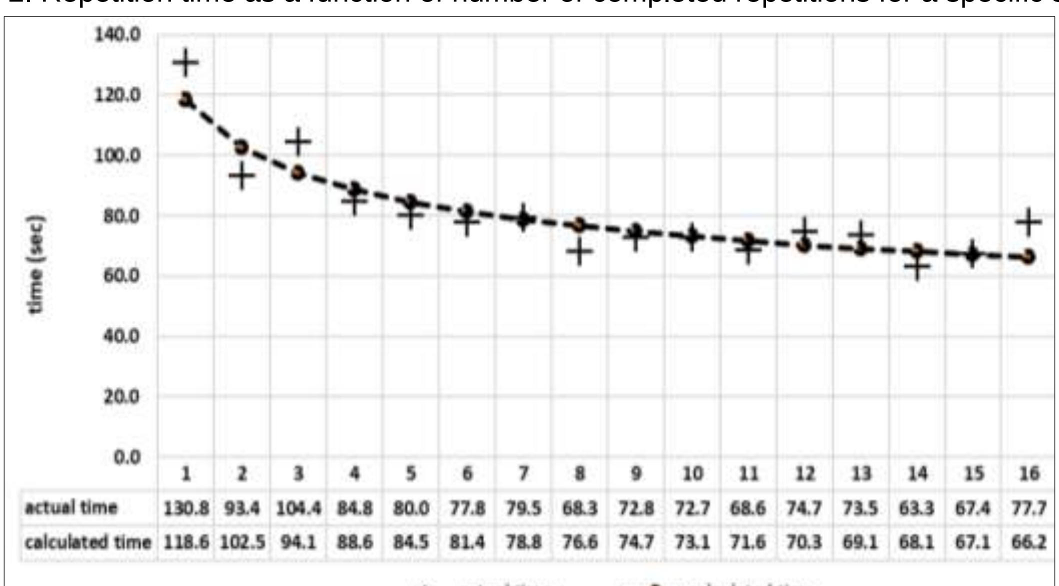
Figure 1. Average actual time for all participants, as a function of repetition number, and the fitted learning curve



As can be traced from figure 1, there is clear evidence for the existence of the learning curve phenomenon. Variation of the data points around the fitted learning curve may be explained, since each participant has a different experience in performing tasks similar to the one used in the study.

Values for each subject had to be calculated to analyze the collected data. Figure 2, presents the data set for a specific subject. This example clarifies variables and parameters used later in this paper.

Figure 2. Repetition time as a function of number of completed repetitions for a specific subject



It was possible to establish the value of the parameters (a, m) using regression analysis, which yielded the best curve fitting. The following are the results obtained from this regression analysis: a=118.5, m=-0.21, R= -0.96

Substituting the value of the parameters into Equation (1), the following equation is obtained to present the data of this participant:

$$(4)\ t(s) = 118.5 \cdot s^{-0.21}$$

Also, substituting the value m= -0.21 into Equation (2), we obtain SL = 0.86. That is, for every doubling of the number of repetitions, time decreases to 86% of the base repetition time. For example, the forecasted time for repetition 32 for this subject is t(32) = t(16)*0.86 = 66.2*0.86 = 56.9

The difference between estimated and actual time

A major objective of this study is to evaluate the ability of subjects to estimate future performance for a non-immediate repetition. This is expressed by the differences between estimated values of repetition 16th given after the second repetition and actual time taken to complete repetition 16, as expressed by equation (5)

(5) d = est(16) – act(16)

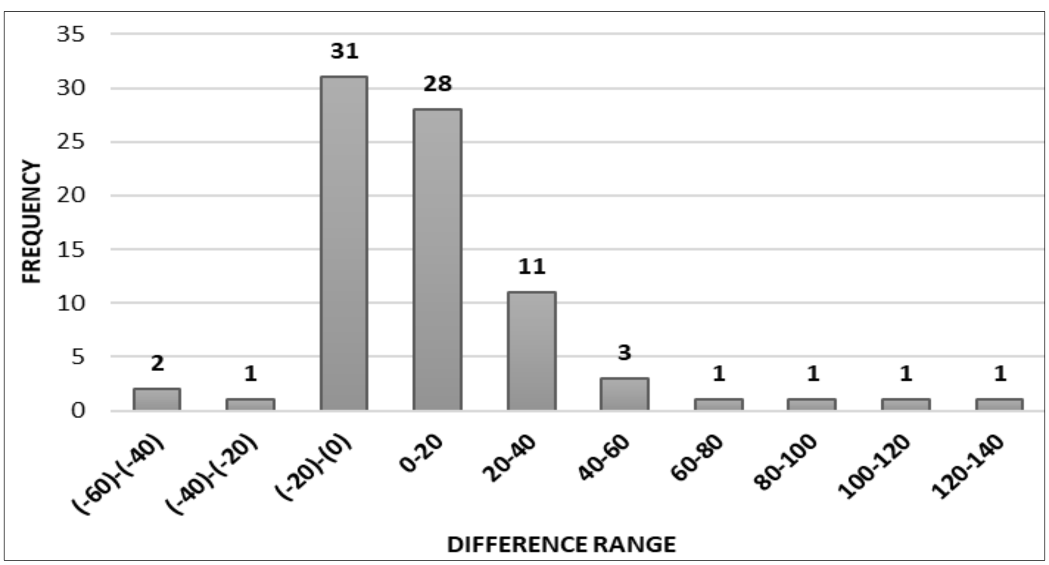
Where:

est – estimated time

act – actual time

Results of the differences for all the participants are demonstrated in figure 3.

Figure 3. Frequency distribution of the differences between estimated and actual performance



The average difference was avg (d)= 9.391 with a standard deviation of std (d)=28.14.

As can be seen from figure 3, the differences are not symmetrically distributed around zero. The hypothesis that the average is significantly different from zero was accepted with a level of confidence of p-value = 0.004. This means that the bias in the ability of subjects to estimate future progress is towards underestimation of their potential improvement. Calculating the average ratio of the difference to actual time overall, the subjects demonstrated a value of 24.3% lower than the estimated value, which is a strong indication of the magnitude of the bias.

Based on figure 2, there is a variation of data points around the learning curve, resulting amongst other reasons, from random disturbances. Therefore, the question is raised whether it is appropriate to use the actual time taken to complete the 16th repetition. A more appropriate measure is the calculated value of t(16), as calculated by its learning curve, since it smooth the random disturbances.

For example, using the data set for the particular subject presented in figure 2, the actual time was act(16)= 77. . Substituting s=16 into Equation 4, the value t(16)= 66.2 is obtained. The difference between the two values is presented by Equation 6.

(6) dif(16)=t(16)-act(16)= -11.5

Where;

dif – is the difference between actual time of the 16th repetition compared to the calculated one based on its learning curve.

The relative difference (rdif) can be calculated by Equation 7.

$$(7)\ \mathrm{rdif}(16) = \frac{\mathrm{t}(16) - \mathrm{act}\,(16)}{\mathrm{act}\,(16)} = -0.148$$

To use t(16) as a reliable representative of act(16), there is a need to verify that there is no significant difference between the two. Thus, rdif(16) for all participants in the study was calculated. Its average was avg(rdif(16))= 0.038 with a standard deviation of std(rdif(16))=0.182. The hypothesis to be tested is that the average of this population is zero. Since the sample size is large enough (that is, 80), a normal distribution may be assumed. The hypothesis was accepted with a level of confidence of 0.05. Therefore, the value of t(16) can be adopted as a proper representative of act(16). The advantage of using the calculated rather than the actual time is due to the smoothing effect that it has. It will also be used below for analyzing actual deviations around the learning curve

Since t(16) can be used to represent act(16), its value for all participants was calculated similarly to the way that it was calculated for the above example.

The ability of subjects to estimate future performance can now be analyzed on the basis of the differences between estimated values of the 16th repetition and the calculated time of repetition 16, as expressed by Equation 8.

(8) dt(16) = est(16) –t(16)

Where:

dt – difference between estimated to calculated time

est – estimated time

t – calculated time

Analysis of all data using the definition of the difference as presented by Equation 8 yielded the following results: The average difference was avg(dt)=8.94 with a standard deviation of std(dt)=27.6.

The parameters of the two distributions of differences, defined by Equations 5 and 8 are summarized in table 1.

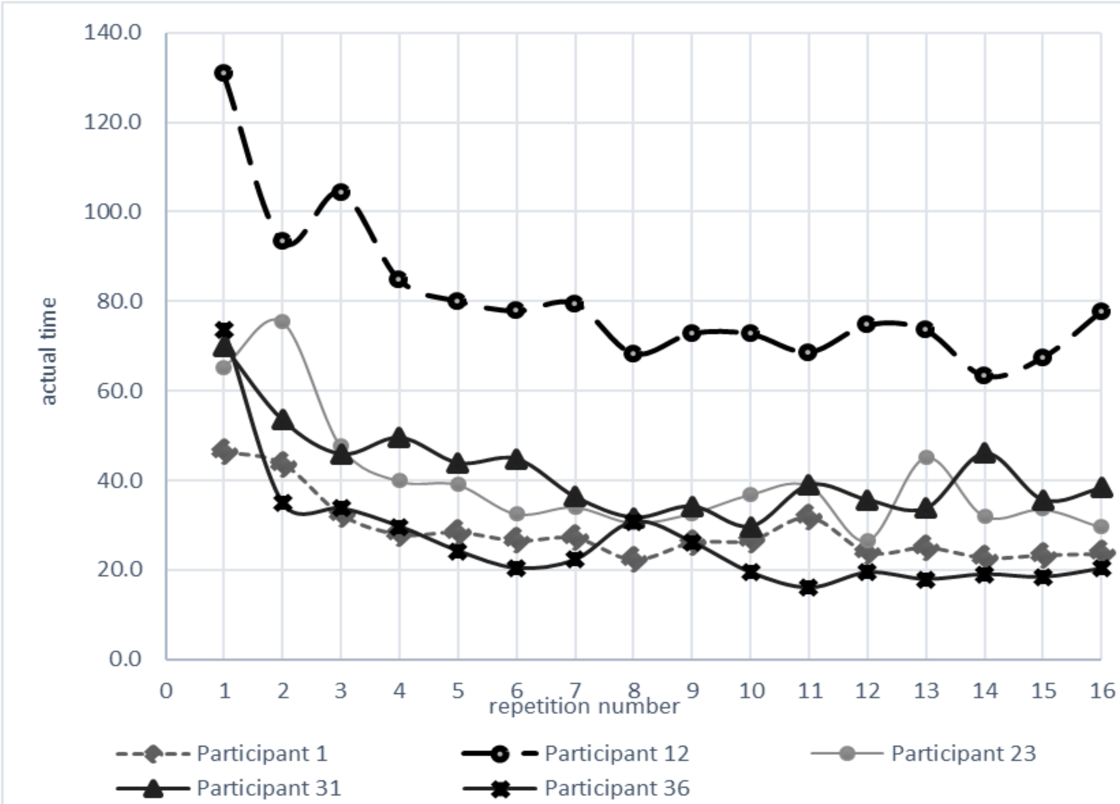
Table 1. Parameters of the two distributions; the differences between estimated and actual time, and the differences between estimated and calculated time

Nature of the difference	Equation	Average	Standard deviation
Estimated – Actual	d = est(16) – act(16)	9.391	28.14
Estimated – Calculated	dt = est(16) –t(16)	8.94	27.6

Comparison of the two distributions via Chi-Square Goodness of Fit Test did not find significant difference between the two distributions, with a 0.05 level of confidence. That is, both may be used for the analysis.

Since each subject has a different experience in performing tasks similar to the one used in the study, the situation is not the same as in the case that all subjects start a task with similar experience. Therefore, the learning curve of each participant may lay in a different portion of the overall learning curve, as depicted for a sample of participants in figure 4.

Figure 4. Learning curves of a sample of 5 participants



As can be traced from figure 4, learning curve of each participant has different parameters due to, among other possible reasons, their previous experience on similar tasks. Previous experience probably has an impact not only on the learning model parameters but also on the deviation around the learning curve. Deviation can be measured by different parameters, such as the "dt" difference defined in Equation 8 or in a proportion as defined by Equation 9.

$$(9)\ \mathrm{pdt}(16) = \frac{\mathrm{dt}(16)}{\mathrm{t}(16)}$$

Where;

pdt(16) expresses the proportion of the difference compared to the number calculated. For example, $pdt(16) = \frac{11.5}{66.2} = 0.174$ for the example analyzed in figure 2.

It makes sense to assume that experience manifests itself not only in average performance but also in the deviation around the average. Therefore, a hypothesis to be tested is that performance deviation around the learning curve decreases, when experience is accrued.

In figure 2 one can observe deviations of the actual data points around the learning curve model for a specific individual. Two ways to present those deviations were chosen for testing the hypothesis. The first one uses the average of the absolute deviation as depicted by Equation 10, and the second uses the average of the relative deviation as can be seen in Equation 11.

$$(10)\ \text{avg(absdev)} = \frac{\sum_{s=1}^{16} ABS(act\,(s) - t(s))}{16}$$

Where;

avg(absdev) – average of the absolute deviations

$$(11)\ avg(relabsdev) = \frac{\sum_{s=1}^{16} \frac{ABS\ (act\ (s) - t(s)}{t(s)}}{16}$$

Where;

avg(relabsdev) – average of the relative absolute deviation.

Table 2 presents example of all the calculations required for the participant whose learning curve is described in figure 2.

Table 2. Example of the calculations required for one participant

s	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
act(s)	130.80	93.40	104.40	84.80	80.00	77.80	79.50	68.30	72.80	72.70	68.60	74.70	73.50	63.30	67.40	77.70
t(s)	118.57	102.49	94.11	88.59	84.53	81.35	78.76	76.58	74.71	73.07	71.62	70.32	69.15	68.08	67.10	66.19
act(s)-t(s)	12.23	-9.09	10.29	-3.79	-4.53	-3.55	0.74	-8.28	-1.91	-0.37	-3.02	4.38	4.35	-4.78	0.30	11.51
abs(act(s)-t(s))	12.23	9.09	10.29	3.79	4.53	3.55	0.74	8.28	1.91	0.37	3.02	4.38	4.35	4.78	0.30	11.51
abs(act(s)-t(s))/t(s)	0.103	0.089	0.109	0.043	0.054	0.044	0.009	0.108	0.026	0.005	0.042	0.062	0.063	0.070	0.004	0.174

The parameters for Equation 4, repeated below, were the result of regression analysis performed on the actual results presented in the row entitled act(s). Then the equation was used to calculate row t(s).

$$(4)\ t(s) = 118.5 \cdot s^{-0.21}$$

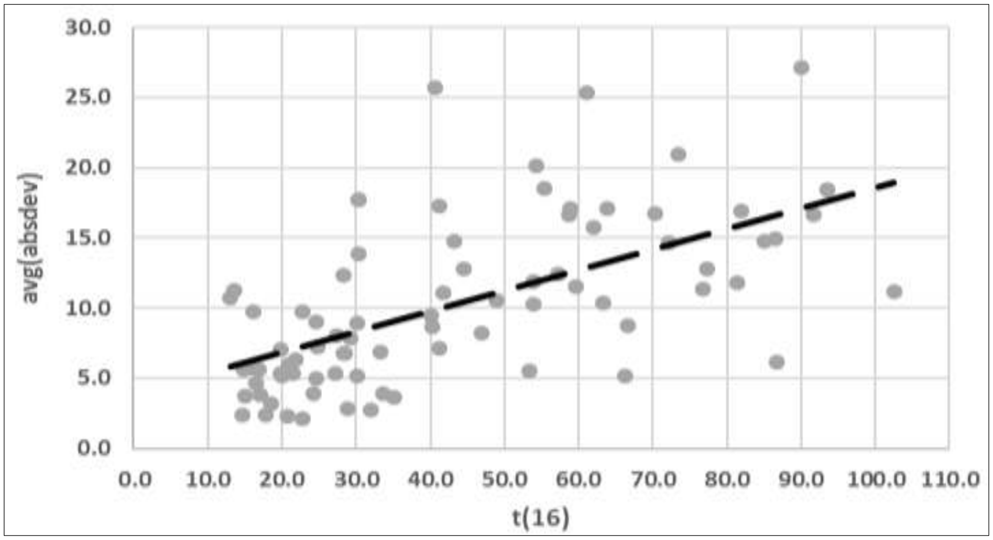
Substituting the values calculated in figure 6 into Equations 10 and 11 obtains:

avg(absdev) = 5.195

avg(relabsdev) = 0.063

The above were the results for the specific participant. The same calculations were performed for all participants. Figure 5 charts the values of the average absolute deviation of all subjects as a function of their final performance as depicted by t(16).

Figure 5. Average absolute deviation of all subjects as a function of their final performance, and as expressed by t(16)



Regression analysis of the data presented in figure 5 yields a correlation coefficient of 0.60 (p=0.001), meaning that there a larger deviation is expected for subjects with less experience.

Repeating similar analysis when deviation is expressed in relative terms (Equation 11) yields a correlation coefficient of -0.31 (p=0.006), meaning that although the absolute deviation of experienced subjects decreases over time, the relative deviation tends to increase.

For further investigation of the relationship between performance level and deviation, participants were divided into two groups based on their final performance as presented by t(16). The median value of t(16) was 34.34 seconds. Therefore, one group (Group I) consists of all participants who obtained lower than the median and the other (Group II) above the median. Statistical results are summarized in table 3.

Table 3. Comparison of the lower to the higher performance groups

Variable to be compared		Group 1 t(16)≤34	Group 2 t(16)≥34	difference of means	p-value
t(16)	Mean	23.14	63.03	-39.89	.000
	Std.deviation	6.12	17.83		
d=est(16)-act(16)	Mean	3.02	15.76	-12.74	.040
	Std.deviation	14.94	36.01		
averabsdev	Mean	6.77	13.75	-6.99	.000
	Std.deviation	3.82	5.51		
avgrelabsdev	Mean	0.224	0.184	0.04	.071
	Std.deviation	0.113	0.075		

Based on the results presented in table 3, the following conclusions can be made: There is a distinctive performance difference between the two groups as demonstrated by t(16); subjects with better performance exhibit better abilities to forecast their performance (see "d" in the figure), as well as having lower deviation of their performance (see averabsdev).

SUMMARY AND CONCLUSIONS

The experiment used for this study was performed with activities typically performed all the time (typing) by all subjects, but with different level of practice. Despite this fact an obvious learning curve was presented, in which the average actual time of the 80 subjects dropped exponentially from 89.2 to 42.6 seconds.

As the subjects probably have different typing experience, it is expected that their learning curve parameters will be significantly different from each other. Results show that actual time for the first repetition was within the wide range of 15-219 seconds, with a learning curve slope ranging from 60% to approximately 100%.

An experienced worker is typically one who has better abilities than an unexperienced worker. Better abilities are expressed by parameters such as performance time and quality of the end product, which was introduced into this experiment as well; a subject was not able to continue to the next repetition unless the previous one was perfectly completed. Therefore, performance time was taken into consideration --- both speed and the quality of work. We further compare the higher performance to the lower performance persons participating in the study. Performance deviation of the lower group is significantly higher, pointing out that experienced workers are characterized not only by higher performance but also by higher stability, as presented by lower performance deviation.

A major objective of the study was to investigate the subjective ability of subjects to estimate future performance. It was found that subjects underestimate their potential improvement by around 24%. That is, if there is a need to forecast performance time of activities such as the one used in this experiment, and subjective forecasting is being, it should be reduced by 24% to compensate for the subjective bias.

A typical organization finds itself quite frequently in situations in which it has to use subjective estimation for forecasting time required to perform activities. Results of this study point out that subjective estimation suffers from consistent bias. Although the magnitude of the bias was found in this study, it is impossible to use it for other activities as long as it has not been validated by using experiments with a different nature of the activities.

Further studies should concentrate on expending experiments of similar nature but of different skills required for performing the operation under study. Repetitions of similar experiments will enable to evaluate if the findings in this research can be applied to other activities as well.

REFERENCES

Brendgen, M., Vitaro, F., Turgeon, L., Poulin, F., & Wanner, B. (2004). Is there a dark side of positive illusions? Overestimation of social competence and subsequent adjustment in aggressive and nonaggressive children. Journal of Abnormal Child Psychology, Vol. 32, P. 305–320.

Buehler, R., Griffin, D., Ross, M. (1994). Exploring the “planning fallacy”: why people underestimate their task completion times. Journal of Personality and Social Psychology, Vol. 67, P. 366–381.

Buehler, R., Griffin, D., Peetz, J. (2010). The planning fallacy: subjective, motivational, and social origins. In P. Z. Mark & M. O. James (Eds.), Advances in experimental social psychology ,43rd ed., P. 1–62.

Chan, A., Hoffmann, E. (2017). Subjective Estimation of Task Time and Task Difficulty of Simple Movement Tasks, Journal of Motor Behavior, Vol. 49, No. 2, P. 185-199

Cole, D. A., Martin, J. M., Peeke, L. A., Serozinski, A. D., & Fier, J. (1999). Children’s over- and underestimation of academic competence: A longitudinal study of gender differences, depression and anxiety. Child Development, Vol. 70, P. 459–473.

König, C., Wirz, A., Thomas K., Weidmann, R. (2015). The Effects of Previous Misestimation of Task Duration on Estimating Future Task Duration. Current Psychology, Vol. 34, No. 1, P. 1–13.

Deliano, M., Tabelow, K., König, R., Polzehl, J. (2016). Improving Accuracy and Temporal Resolution of Learning Curve Estimation for within- and across-Session Analysis, PLoS ONE 11(6), https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0157355

Christoph H. Glock, Eric H. Grosse, Mohamad Y. Jaber, Timothy L. Smunt. (2018). Applications of learning curves in production and operations management: A systematic literature review. Computers & Industrial Engineering, https://doi.org/10.1016/j.cie.2018.10.030

Griffiths, T., Tenenbaum, J. (2011). Predicting the Future as Bayesian Inference: People Combine Prior Knowledge With Observations When Estimating Duration and Extent, Journal of Experimental Psychology, Vol. 140, No. 4, P. 725–743.

Grosse, E., Glock, C., Müller, S.,. (2015). Production economics and the learning curve: A meta-analysis. International journal production economics, Vol. 170, P. 401–412.

Kahneman, D., & Tversky, A. (1979). Intuitive prediction: Biases and corrective procedures. TIMS Studies in the Management Sciences, Vol. 12, P. 313–327.

Katahira, K., Yuki, S., Okanoya, K., (2017). Model-based estimation of subjective values using choice tasks with probabilistic feedback. Journal of Mathematical Psychology, Vol. 79, P. 29–43.

Kruger, J., Evans, M. (2004). If you don’t want to be late, enumerate: unpacking reduces the planning fallacy. Journal of Experimental Social Psychology, Vol. 40, P. 586–598.

Kazi, S., Samartzi, S., Koustoumbardis, M. (2014). Duration estimation of writing: The influence of emotionality on subjective time. Social and Behavioral Sciences, Vol. 126, P. 249 – 250.

Rodon, C., Meyer, T. (2012). Searching information on the web and planning fallacy: a pilot investigation of pessimistic forecasts. European Review of Applied Psychology, Vol. 62, P. 103-109.

Roy, M., Mitten, S., Christenfeld, N. (2008). Correcting memory improves accuracy of predicted task duration. Journal of Experimental Psychology. Applied, Vol. 14, P. 266–275.

Thomas, K., Handley, S. (2008). Anchoring in Time Estimation. Acta Psychologica, Vol.127, No 1, P.24-29.

Wright, S. S. (2000). Looking at the self in a rose-colored mirror: Unrealistically positive self-views and academic performance. Journal of Social and Clinical Psychology, Vol. 19, P. 451–462. http://dx.doi.org/10.1521/jscp.2000.19.4.451.