

# An Estimation of the Learning Curve Effect on Project Duration with Monte Carlo Simulation

Levente Mályusz<sup>1\*</sup>, Anita Varga<sup>2</sup>

<sup>1</sup> Department of Construction Technology and Management, Faculty of Architecture, Budapest University of Technology and Economics, 1111 Budapest, Műegyetem rkp. 3, Hungary

<sup>2</sup> Department of Differential Equations, Faculty of Natural Sciences, Budapest University of Technology and Economics, 1111 Budapest, Műegyetem rkp. 3, Hungary

\* Corresponding author, e-mail: [lmalyusz@ekt.bme.hu](mailto:lmalyusz@ekt.bme.hu)

Received: 27 June 2018, Accepted: 27 June 2018, Published online: 02 July 2018

## Abstract

The aim of this paper is to estimate learning curve effect on project duration with the mean of project scheduling techniques. To measure this effect only one assumption is taken: the activity time individuals / groups take to perform an activity decreases at a given rate as experience is gained with the activity. Unfortunately this effect directly is not taken into account by project management software. In some software after scheduling, supervisor manually can switch on the "as soon as possible" or "as late as possible" buttons on an activity.

Monte Carlo simulation was used to model the risks in the total project durations. It is assumed that the (normal) durations of the activities can vary according to the beta distribution. The minimum estimate is 95 % of the original (normal) duration, and the maximum estimate is 140 % of the original (normal) duration. We assumed that most likely value is the (normal) duration of each activity. The learning effect on project duration with the help of test problems and real problems was investigated. In test problems learning effect can occur between two consecutive activities. These pairs are chosen randomly. After calculating project duration, these pairs are allocated to be closer to each other using the predecessor's total float time. It is assumed that the duration of impending repetitive activities is shorter due to the learning curve effect if the gap between consecutive activities is small enough. This iteration is carried out until it is not possible to shorten the successor's activity time in a pair. It is shown that this effect brings a 2-3 % shorter project duration meanwhile variance is also left in a 1-2 % range. Numerical tests were implemented by XPRESS-Mosel Optimization Software.

## Keywords

Project Scheduling, Learning Effect, Learning Curve, integer programming

## 1 Introduction

In the literature some researches, depending on at which level the phenomenon occurs (individual, group, firm, industry), distinguish learning, progress and experience curves for the same phenomenon. We use the term learning curve to encompass the term "progress curve" and "experience curve" also.

In this paper, project scheduling in mathematical terms means finding the longest path in a directed graph, where vertices and directed edges are given. Also, there is a given integer number assigned to each edge. In engineering terms, directed edges represent activities or connections between activities, vertices are nodes or events, and integer numbers represent activity times or time lags between activities. It is assumed that the learning effect

can occur and result in a smaller activity time of a given activity if the same group of workers perform a similar activity as an immediate predecessor of the given activity. The question is: what is or what can be the cumulative effect of the reductions of activity times in the project scheduling network. This effect is calculated using calendar days, which leads to a more complex mathematical model and algorithm than calculating using only working days (Mályusz and Varga, 2016).

In a construction project, the general contractor distributes the work among subcontractors. Normally general contractor organizes its work based on consecutive technology steps. In this way the subcontractors often should do their activities interrupted.

It is obvious that they can reduce their costs if they work continuously. There are two main reasons for this: first, they can reduce their construction costs; second, the work will be completed sooner because of the learning effect. Unfortunately, in the early phase of scheduling, this effect is not considered and not supported by project management softwares.

In a multi-project environment, the learning see (Wu and Sun, 2006) effect of staff was considered when periodically scheduling the tasks for each project and assigning staff to the tasks. The solution leads to a mixed nonlinear program for project scheduling and staff allocation problems, which considers the learning effect of staff. A genetic algorithm (GA) is proposed to solve the problem. (Zha and Zhang, 2014) investigate the project scheduling problem with multiskill learning effect, where both autonomous and induced learning is considered.

In practice, project scheduling methods suffer from a lack of precision; consequently, it is a significant challenge to create a realistic and usable project schedule. It is difficult and time-consuming to estimate time, assign resources, determine interdependencies between tasks, and manage changes. It is, therefore, important to identify and investigate the differences between the practice and theory of scheduling methods (Francis et al., 2013).

In construction project management, the appropriate scheduling of a project is an essential problem. Estimation of an activity's time is a crucial part of the schedule. There is little information in the literature about the use of learning curves in scheduling, although it seems that the principle of learning curves is gathering ground in the scheduling of repetitive construction operations (Hinze and Olbina, 2009; Zahran et al., 2016). In (Hajdu, 2015), the learning curve effect on linear scheduling method is discussed. However, it should be noted that the impact of learning curves is not calculated in recent management software (Fini et al., 2016).

## 2 Learning Curve

Psychologist researchers at the turn of the 19th century focused on behavior of individuals. They found that time individuals took to perform a task and the number of errors they made decreased at a certain rate after repetitions see (Ebbinghaus, 1964; Thorndike, 1898; Thurstone, 1919).

Researchers also found that errors made by groups taking repetitive activities also are decreased at a certain level as groups gained experiences (Guetzkow and Simon, 1955; Leavitt, 1951).

It is a debate among researchers whether organizational learning is a consequences of changes in behavior or changes in cognitions. It is obvious that when firm or group does repetitive works, members might learn who is good at what, and how they organize their work better. Presumably in construction industry the following two changes have bigger effects: members learn how they layout working site and how they can fit their current work to the local regulations and conditions.

It is an open question and further research is needed in how this phenomena reflects in changes of project scheduling network. The learning curve formulation:

$$y = a_0 + (a - a_0)(x + B)^b .$$

In the learning curve formulation the standard measure of experience is the cumulative number of units produced. So the measure is calculated by summing the total number of products from the start through the end of each time period. The cycle number is denoted by  $x$ ,  $y$  is the time required to complete cycle  $x$  in labor hour per output,  $a$  is the time required to complete the first cycle,  $a_0$  is the minimum required time complete a cycle,  $b$  is a learning coefficient.  $B$  expresses the number of units produced before the first unit, so it is an experience factor. The value of  $B$  will be in the range of 0-10 (Gottlieb and Haugbølle (Hasanzadeh et al., 2016; Kara and Kayis, 2005: 209). Here  $y$  can represent not only time or cost but a wide range of outcomes of production for instance: defects per unit, or accidents per unit (Greenberg, 1971).

When  $B$  and  $a_0$  are 0, then we get back the original Wright's formula.

$$\ln y = \ln a + b \ln x; \vee y = ax^b = ax^{\log_2 r} , \quad (1)$$

where  $r$  is the rate of learning.

Wright discovered that when the production / cycles doubles the cumulative labor time / cost decreases at a constant rate, that is, the learning rate. So learning rate is the constant rate with which cumulative labor time/cost decreases when the production / cycles doubles in a linear  $\log x$ ,  $\log y$  model. This feature of the learning rate comes from the logarithms nature and true only in linear  $\log x$ ,  $\log y$  model.

In the literature, there are several papers on learning effect of construction: (Oglesby et al., 1993; Drewin, 1982; Teplitz, 1991; Everett and Farghal, 1994; 1997; Lutz et al., 1994; Lam et al., 2001; Couto and Texiera, 2005). Learning curve theory can be applied to predict cost and time, generally in units of time, to complete repetitive activities (Malyusz and Pem, 2014; Hasanzadeh et al., 2016).

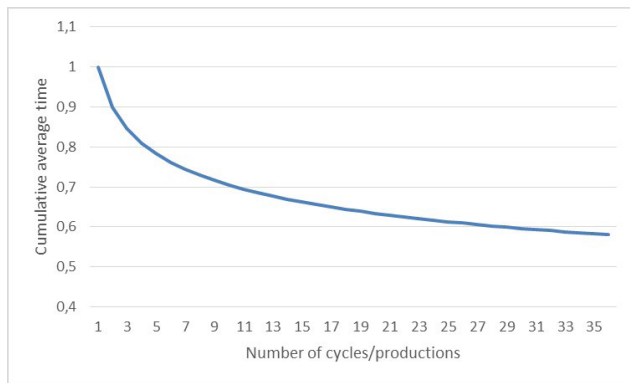


Fig. 1 Learning curve with 90 % learning rate

Several researchers have suggested that Wright's model is the best model available for describing the future performance of repetitive work (Everett and Farghal, 1994; Couto and Texiera, 2005). In the exponential average method (Mályusz and Pém, 2013),  $\alpha = 0.5$  yielded the most accurate predictions. Of course, there is no consensus on which model provides the best fit and predictability for construction data (Srour et al., 2016). Consequently, more theoretical and experimental investigations are necessary to adjust a model according to the real problems.

In the construction industry, the learning rate is between 85-95 %. According to the practice and theory as well after a certain amount of produced "unit" there is no significant reduction in time and cost.

Using an example of 90 % of the learning rate, if a job is ten days, a repetition of that is eight days if the working conditions are similar. The learning curve effect does not always apply, of course. It flourishes where certain conditions are present; it is also necessary for the process to be a repetitive one. Additionally, there needs to be a continuity of workers without any abrupt stops during the production process. When the learning curve effect, on occasions, comes to an abrupt stop, graphically, the curve jumps up (Ferivanto et al., 2015).

### 3 Project Scheduling Model

In this paper, the concept of activity on node network is followed. The relationships between the activities can be represented by a  $G = (V, E)$  directed graph, where  $V$  is the set of nodes ( $|V| = n$ ), and  $E: V \rightarrow V$  denotes the set of directed edges. Each node corresponds to an activity, and the relations are represented by the directed edges.

In the case of the exponential time algorithm, the relations can be both maximal and minimal type, but the heuristic algorithm is only able to handle minimal type

finish-start relations. Although, there are four relationships that can be defined between activities, for the sake of simplicity, from now on only the Finish Start relationship is used and its variants, namely: FS $k$ , max FS1, max FS0. The maximal precedence relationship describes the maximum allowable time between the start / finish point of the preceding and the start / finish point of the succeeding activity (in calendar days).

A  $w_e$  weight is assigned to each  $e \in E$  directed edge, it determines the length of the relation given by the  $e$  directed edge (in calendar days). It is also assumed that  $G$  is acyclic.

The calendar vector  $\mathbf{c}$  is given in advance (It is determined by according to the time period what is used during the calculations).  $\mathbf{c}$  is a binary vector,  $c(i) = 1$ , if day  $i$  is a working day, and 0 otherwise. It is supposed that  $\mathbf{c}$  is longer than the maximal possible total project duration.

Two positive integer variables are assigned to each node  $v$  (the starting and finishing time of the corresponding activity):  $x_{2v-1}$  and  $x_{2v}$ . A predefined  $d_v$  positive integer number is also ordered to each  $v$  node (the duration of the corresponding activity).

To satisfy the precedence constraints defined by the directed edges and the weights, for all  $e = (i, j) \in E$  directed edge, the following constraints is added:

$$x_{2j-1} - x_{2i} \geq w_e.$$

To model the learning curve effect, it is supposed that there are special learning relations (edges).  $LR$  denotes the set of these edges, and  $K$  denotes the number of these relations:

$$LR \subseteq E, |LR| = K.$$

To make calculations easier, it is also supposed that the duration of the endpoints of the learning edges is equal. For ease of understanding, in this heading, it is supposed that  $d_j = 10$ , if  $j$  is an endpoint of a learning edge, and it can be shortened to 9 or 8 days. It can be rescaled, according to the learning curve defined in the previous heading.

If node  $j$  is not an endpoint of a learning edge (so its duration cannot be shortened by the learning curve effect), the following constraint is added:

$$\sum_{k=x_{2j-1}}^{x_{2j}} c_k = d_j.$$

For a learning relation  $(i, j) \in LR$ , the following constraints are added to the model to ensure the shortening of the duration of activity  $j$  due to the learning curve effect:

If there is no working day between the two activities (so the sum of the calendar vector between the two indices is 0), then the length of the second activity is eight days:

$$\text{If } \sum_{k=x_{2i}}^{x_{2j}-1} c_k = 0, \text{ then } x_{2j} - x_{2j-1} = 8.$$

If one working day passes between the finishing time of  $i$  and the starting time of  $j$ , then the length is nine days:

$$\text{If } \sum_{k=x_{2i}}^{x_{2j}-1} c_k = 1, \text{ then } x_{2j} - x_{2j-1} = 9.$$

If there are at least two working days between the activities, then the duration of the second one is ten days:

$$\text{If } \sum_{k=x_{2i}}^{x_{2j}-1} c_k \geq 2, \text{ then } x_{2j} - x_{2j-1} = 10.$$

Our objective function is: (It is equivalent to minimizing the total project duration)

$$\min x_{2n}.$$

To sum up, mathematical model is the following:

Given:  $G = (V, E)$ , acyclic directed graph,  $w_e$  integer number to each  $e \in E$  directed edge,  $c$  is a binary vector,  $c(i) = 1$ , if day  $i$  is a working day, and 0 otherwise,  $d_v$  positive integer number is also ordered to each  $v$  node.

Find: positive numbers  $x_{2v-1}$  and  $x_{2v}$  assigned to each node  $v$ .

$$\min x_{2n} \quad (2)$$

s. t.

$$\forall (i, j) \in (E \setminus LR): \quad x_{2j-1} - x_{2i} \geq w_e \quad (3)$$

$$\forall j: \exists i: (i, j) \in LR: \quad \sum_{k=x_{2i}}^{x_{2j}-1} c_k = d_j \quad (4)$$

$$\forall (i, j) \in LR: \quad \text{If } \sum_{k=x_{2i}}^{x_{2j}-1} c_k = 0, \text{ then } x_{2j} - x_{2j-1} = 8. \quad (5)$$

$$\text{If } \sum_{k=x_{2i}}^{x_{2j}-1} c_k = 1, \text{ then } x_{2j} - x_{2j-1} = 9. \quad (6)$$

$$\text{If } \sum_{k=x_{2i}}^{x_{2j}-1} c_k \geq 2, \text{ then } x_{2j} - x_{2j-1} = 10. \quad (7)$$

$$x_i > 0, \quad x_i \text{ integer } \forall i = 1, 2, \dots, 2n.$$

This integer programming model is similar to the one is presented in (Mályusz and Varga, 2016), for the working

days calculation. In that case, the constraints defining the learning relations can be linearized. In this model, unfortunately, it is not possible.

#### 4 Numerical Test

To examine the learning curve effect on artificial projects and to be able to compare the exact results with those from the heuristic algorithm, different test problems were randomly generated.

It is assumed that the activities could be ordered to equal length rows; the number of rows and columns are denoted by  $M$  and  $S$ , respectively. A starting and finishing activities are added, the first elements in each row are connected with the starting activity, and the last elements are connected with the finishing activity. Therefore, each graph contains  $M \cdot S + 2$  nodes. The adjacent nodes in each row are connected with finish-start relations. Between different rows, edges (connections) were added,  $K$  denotes the number of the learning edges, and  $L$  is the number of edges between different rows without learning effect. An example can be seen in Fig. 2. The learning edges are highlighted in red.

To examine the changes in the total project durations, we generated 2500 problem instances according to the beta distribution, based on an example project. Our example project consists of 102 activities and it is generated with the same method that we used in (Wright, 1936). The project contains five pairs of learning edges; it means that we have five pairs of repetitive activities where the second activity's duration can be shortened. Using the integer programming model from (Wright, 1936), average of project duration of the example project is 238 days without taking into consideration the learning curve effect, and it can be shortened to 232 days using learning effect meanwhile variance went from 2.5 to 2.7 days.

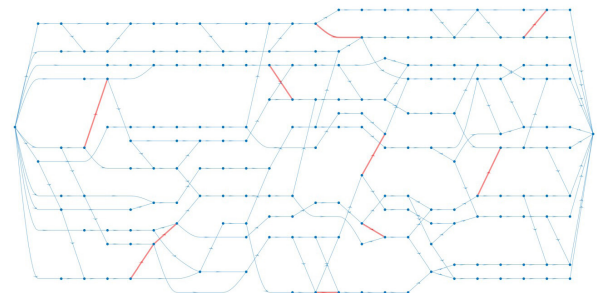


Fig. 2 A randomly generated network with 202 activities and 10 learning connections

The results of the simulation can be seen in Fig. 1. The horizontal axis shows the total project durations, and the height of the column describes the frequency of each total project duration in the Monte Carlo simulation. The blue columns show the results in case of the integer programming model that takes into consideration the learning curve effect (Wright, 1936), while the red columns at the background illustrates the results without the learning curve effect.

## 5 Concluding Remarks

In this paper learning effect on project duration is investigated with a help of Monte Carlo simulation. It is presented that the cumulative learning effects of activities in a project can cause a 2-3 % reduction in project duration, meanwhile variance is about 1 % of project duration. Since the exponential time algorithm is very slow for real problems, further development of the heuristic algorithm is necessary. An investigation of learning effect

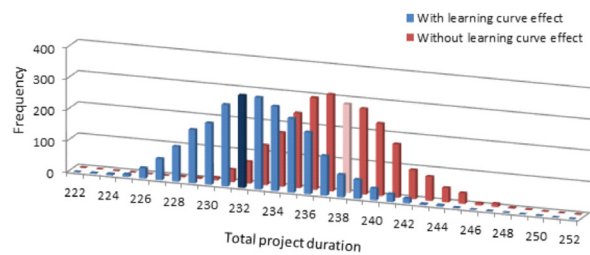


Fig. 3 Changes in the total project duration

on project cost is also an interesting research topic and deserves further consideration.

It is an open question and further research is needed in how the following phenomenon reflects in changes of project scheduling network of consecutive projects: how labors learn how they layout working site, how they can fit their current work to the local regulations and conditions and how they organize their work better.

## References

- Couto, J. P., Texiera, J. C. (2005) "Using linear model for learning curve effect on highrise floor construction", *Construction. Management Economics*, 23(4), pp. 355–364.  
<https://doi.org/10.1080/01446190500040505>
- Drewin, F. J. (1982) "Construction productivity: Measurement and improvement through work study", Elsevier Science, New York.
- Ebbinghaus, H. (1964) "Memory: A contribution to experimental psychology", Dover, New York.
- Everett, G., Farghal, H. (1994) "Learning Curve Predictors for Construction Field Operations", *Journal of Construction Engineering and Management*, 120(3), pp. 603–616.  
[https://doi.org/10.1061/\(ASCE\)0733-9364\(1994\)120:3\(603\)](https://doi.org/10.1061/(ASCE)0733-9364(1994)120:3(603))
- Everett, G., Farghal, H. (1997) "Data representation for predicting performance with learning curves", *Journal of Construction Engineering and Management*, 123(1), pp. 46–52.  
[https://doi.org/10.1061/\(ASCE\)0733-9364\(1997\)123:1\(46\)](https://doi.org/10.1061/(ASCE)0733-9364(1997)123:1(46))
- Feriyanto, N., Saleh, C., Badri, H. M., Deros, B. M., Pratama, Y. (2015) "Implementation learning and forgetting curve to predict needs and decrease of labor performance after break", *Jurnal Teknologi*, 77(27), pp. 135–140.  
<https://doi.org/10.11113/jt.v77.6909>
- Fini, A. A. F., Rashidi, T., Akbarnezhad, A., Travis Waller, S. (2016) "Incorporating Multiskilling and Learning in the Optimization of Crew Composition", *Journal of Construction Engineering and Management*, 142(5).  
[https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001085](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001085)
- Francis, A., Bibai, J., Miresco, E. T. (2013) "Simulation of scheduling logic using dynamic functions", *Management, Procurement and Law*, 166(3), pp. 145–158.  
<https://doi.org/10.1680/mpal.11.00049>
- Gottlieb, S. C., Haugbølle, K. "The Repetition Effect in Building and Construction Works, A Literature Review", Danish Building Research Institute, Aalborg University, 2010. [online] Available at: <https://sbi.dk/Assets/The-repetition-effect-in-building-and-construction-works/sbi-2010-03-pdf.pdf> [Accessed: 26 June 2018]
- Greenberg, L. (1971) "Why the Mine Injury Picture is out of Focus", *Mining Engineering*, 23, pp. 51–53.
- Guetzkow, H., Simon, H. A. (1955) "The impact of certain communication nets upon organization and performance in task-oriented groups", *Management Science*, 1(3-4), pp. 233–250.  
<https://doi.org/10.1287/mnsc.1.3-4.233>
- Hajdu, M. (2015) "Continuous Precedence Relations for Better Modelling Overlapping Activities", *Procedia Engineering*, 123, pp. 216–223.  
<https://doi.org/10.1016/j.proeng.2015.10.080>
- Hasanzadeh, M. R., Shirani, B. A., Raissi Ardali, G. A. (2016) "An Improved Performance Measurement Approach for Knowledge-Based Companies Using Kalman Filter Forecasting Method", *Mathematical Problems in Engineering*, 2016.  
<https://doi.org/10.1155/2016/4831867>
- Hinze, J., Olbina, S. (2009) "Empirical analysis of the learning curve principle in prestressed concrete piles", *Journal of Construction Engineering and Management*, 135(5), pp. 425–431.  
[https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000004](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000004)
- Kara, S., Kayis, B. (2005) "The Effect of the Learning Process in Concurrent Engineering Projects", *Concurrent Engineering*, 13(3), pp. 209–217.  
<https://doi.org/10.1177/1063293X05056473>
- Lam, K. C., Lee, D., Hu, T. (2001) "Understanding the effect of the learning-forgetting phenomenon to duration of project construction", *International Journal of Project Management*, 19(7), pp. 411–420.  
[https://doi.org/10.1016/S0263-7863\(00\)00025-9](https://doi.org/10.1016/S0263-7863(00)00025-9)

- Leavitt, H. J. (1951) "Some effect of certain communication patterns on group performance", *Journal of Abnormal and Social Psychology*, 46(1), pp. 38–50.  
<https://doi.org/10.1037/h0057189>
- Lutz, J. D., Halpin, D. W., Wilson, J. R. (1994) "Simulation of learning development in repetitive construction", *Journal of Construction Engineering and Management*, 120(4), pp. 753–773.
- Mályusz, L. (2016) "Learning Curve Effect on Project Scheduling", *Procedia Engineering*, 164, pp. 90–97.  
<https://doi.org/10.1016/j.proeng.2016.11.596>
- Mályusz, L., Pém, A. (2013) "Prediction of the Learning curve in Roof Insulation", *Automatization in Construction*, 36, pp. 191–195.  
<https://doi.org/10.1016/j.autcon.2013.04.004>
- Mályusz, L., Pém, A. (2014) "Predicting future performance by learning curves", *Procedia-Social and Behavioral Sciences*, 119, pp. 368–376.  
<https://doi.org/10.1016/j.sbspro.2014.03.042>
- Mályusz, L., Varga, A. (2016) "An Estimation of the Learning Curve Effect on Project Scheduling with Working Days Calculation", *Periodica Polytechnica Architecture*, 47(2), pp. 104–109.  
<https://doi.org/10.3311/PPar.10387>
- Oglesby, C. H., Parker, H. W., Howell, G. A. (1993) "Productivity improvement in construction", McGraw-Hill, New York, 1989.
- Srour, J. F., Kiomjian, D., Srour, I. M. (2016) "Learning Curves in Construction: A Critical Review and New Model", *Journal of Construction Engineering and Management*, 142(4).  
[https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001096](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001096)
- Teplitz, C. J. (1991) "The Learning Curve Deskbook: A Reference Guide to Theory, Calculations, and Applications", Quorum Books, Westport, Conn.
- Thorndike, E. L. (1898). "Animal intelligence: An experimental study of the associative processes in animals", *The Psychological Review: Monograph Supplements*, 2(4), pp. 1–109.  
<https://doi.org/10.1037/h0092987>
- Thurstone L. L. (1919) "The learning curve equation", Kessinger Whitefish, MT, USA.
- Wright, T. P. (1936) "Factors affecting the cost of airplanes", *Journal of Aeronautical Science*, 3(4), pp. 122–128.  
<https://doi.org/10.2514/8.155>
- Wu, M. C., Sun, S.-H. (2006) "A project scheduling and staff assignment model considering learning effect", *The International Journal of Advanced Manufacturing Technology*, 28(11-12), pp. 1190–1195.  
<https://doi.org/10.1007/s00170-004-2465-0>
- Zahran, K., Nour, M., Hosny, O. (2016) "The Effect of Learning on Line of Balance Scheduling: Obstacles and Potentials", *International Journal of Engineering Science and Computing*, 6(4), pp. 3831–3841.  
<https://doi.org/10.4010/2016.889>
- Zha, H., Zhang, L. (2014) "Scheduling Projects with Multiskill Learning Effect", *The Scientific World Journal*, 2014.  
<https://doi.org/10.1155/2014/731081>