

**ALICJA GRZEŚKOWIAK**  
Wrocław University of Economics, Poland

## **SELF-LEARNING AS A METHOD OF IMPROVING SKILLS BY ADULT POLES - EVIDENCE FROM NON-METRIC SURVEY DATA**

### **Abstract:**

Nowadays, in a rapidly changing environment, it is essential to participate in lifelong learning activities in order to update and develop knowledge and qualifications. There are various forms of learning from highly formalized and structuralized to informal ones which allow to improve knowledge and skills. European Union educational policy underlines the necessity of lifelong learning involvement. Unfortunately the participation rate in continuous education in Poland is far beyond European benchmarks. The official statistics concern mostly the engagement in formal and non-formal educational processes. This paper is focused on self-learning activities of adult Poles. The analyses are carried out on the basis of data from a nationwide survey. As the data are categorical adequate statistical methods are applied to reveal and describe the regularities concerning this aspect of continuous education.

### **Keywords:**

longlife learning, informal learning, non-metric data

**JEL Classification:** I21, J24

## 1 Introduction

Participation in lifelong learning is one of the objectives of European educational strategy included in Education and Training 2010 and Education and Training 2020 programmes. Continuous education becomes necessary in the quickly changing environment. There are various forms of learning which are covered by Eurostat into three categories (Classification of learning activities - Manual 2006, p.13):

- formal, which takes place in the educational institutions like schools, colleges and universities and is a part of typical educational "ladder",
- non-formal which is organized and structuralized but it is not a part of the formal system,
- informal, which is "...intentional, but it is less organized and less structured .... and may include for example learning events (activities) that occur in the family, in the work place, and in the daily life of every person, on a self-directed, family-directed or socially directed basis" (*Ibidem*, p. 13)

The formal and non-formal education are broadly depicted in official statistics while less attention is paid to the third category, i.e. informal learning. Up-to-date information on lifelong learning in European Union is available through Labour Force Survey (LFS). A detailed insight is available due to Adult Education Survey. The educational data collected within surveys are used, among others, to construct one of the key education indicators ("participation of people aged 25-64 in education and training") as well as many others which are widely applied in international comparisons and policy monitoring (see e.g., *Adults in Formal Education ... 2011, Progress Towards the Common European Objectives in Education and Training... 2011*) and when considering the situation at the national level (see, e.g. Grześkowiak 2013, Kryńska (ed.) 2013, *Lifelong learning ... 2000*). However, the key lifelong learning indicator reflects the involvement either in formal or non-formal education, so additional information must be used to evaluate the scope and the importance of the third category of learning. Hence, this study is focused on informal education and is based on information on self-learning available from a nationwide survey on human capital in Poland.

The main objective of the study is the analysis of the self-learning activities of Poles on the basis of non-metric survey data. Two specific objectives are formulated as follows:

- evaluation of interdependence between self-learning declaration and various socio-demographic factors,
- assessment of co-occurrences of various self-learning activities.

As the dataset comprises categorical variables, adequate analytical approaches are applied to achieve the defined objectives.

## 2 Data characteristics and analytical methods

The study is carried out on the basis of data from the third wave of the survey on human capital which was conducted in Poland in 2012 (N = 17 600). The respondents in productive age answered the question if in the past 12 months they learned something new or gained experience by various types of self-learning activities (*Bilans kapitału ludzkiego w Polsce. Badania ludności 2012*, p. 24):

- activities supported by family members,
- activities supported by friends or colleagues,
- activities based on books, journals and other printed materials,
- activities based on computer programs and Internet,
- activities based on programmes on TV or radio,
- activities in museums, exhibitions, galleries and science centers,
- other.

One of the objectives of the study is to evaluate the influence of chosen socio-demographic factors on the participation in informal learning. The factors in question are as follows:

- gender (dichotomous variable: men, women),
- age (five categories: 18-24, 25-34, 35-44, 45-54, 55-64),
- education level (four categories: lower secondary or less, vocational, upper secondary, tertiary),
- place of residence (three categories: rural area, smaller town – to 99 000 inhabitants, bigger town/city – more than 99 000 inhabitants).

Two statistical techniques appropriate for non-metric variables are applied to assess associations between the involvement in self-learning and aforementioned factors. As this is a bivariate analysis of two nominal or ordinal variables, chi-square statistic is used to verify if there are significant dependencies between the declaration of self-learning and socio-demographic characteristics of respondents. However, chi-square statistics cannot be treated as a measure of the strength of association since it depends of the number of observations. Hence, the contingency coefficient is applied to evaluate and compare the dependencies. This coefficient proposed by Pearson takes values between 0 and 1 and is calculated according to the formula (Everitt 1992, p. 57):

$$P = \sqrt{\frac{\frac{\chi^2}{N}}{1 + \frac{\chi^2}{N}}}$$

The contingency coefficient equal to 0 occurs in case of independence and higher values indicate a stronger association between two categorical variables.

The other specific objective of the study is focused on the assessment of co-occurrences of various self-learning activities. As the set of the activities comprises many elements a multivariate approach is indispensable in order to analyze the interdependencies simultaneously. Some multivariate methods are designed to deal with non-metric data, namely correspondence analysis (see, e.g., Greenacre 2010) or log-linear modeling (see, e.g., Christensen 1997). Although widely used, they have certain limitations due to visualization problems or sparse matrixes especially in case of a high number of variables. For these reasons another analytical approach selected from unsupervised learning methods is chosen to search for relationships among the responses. This approach based on association rules mining gives opportunity to reveal interesting relationships in large datasets (Cios et al. 2007, p.289). An association rule is given as an implication  $A \Rightarrow B$ . In this research it can be expressed as a statement "if self-learning activity A then self-learning activity B". Detection of the association rules allows to find categories often mentioned together by respondents. The aim of the unsupervised learning algorithm is to find association rules which fulfill some quality criteria determined by the researcher. The most common evaluation concerns the strength and the certainty of a rule and is performed by support and confidence measures (Cios et al. 2007, p. 290). The support measure represents the probability of joint occurrence of statements A and B and confidence indicator is defined as the conditional probability of B given A (Larose 2013, p. 187-189). As these probabilities are not known they are estimated on the base of the given sample. For convenience reasons they are often expressed as percentages. In this study as acceptable association rules are considered those who satisfy the following minimum thresholds: support = 10% and confidence = 10%.

### **3 Socio-demographic determinants of self-learning**

Surveys on human capital in Poland were conducted in 2010, 2011 and in 2012 giving an opportunity to compare the attitudes to improving competences by adults in productive age. According to Szczucka, Turek & Worek (2012, p. 29) 36% of Poles declared that they had improved competences in any form in last 12 months but only 16% of respondents admitted that they had carried out self-learning activities (2012 survey). However, the participation rate in informal education tends to incline which is a positive sign. The difference between 2010 and 2012 self-learning rate is six percentage points (Szczucka, Turek & Worek 2012, p.29).

The involvement in informal learning with respect to chosen socio-demographic factors is presented in Figure 1.

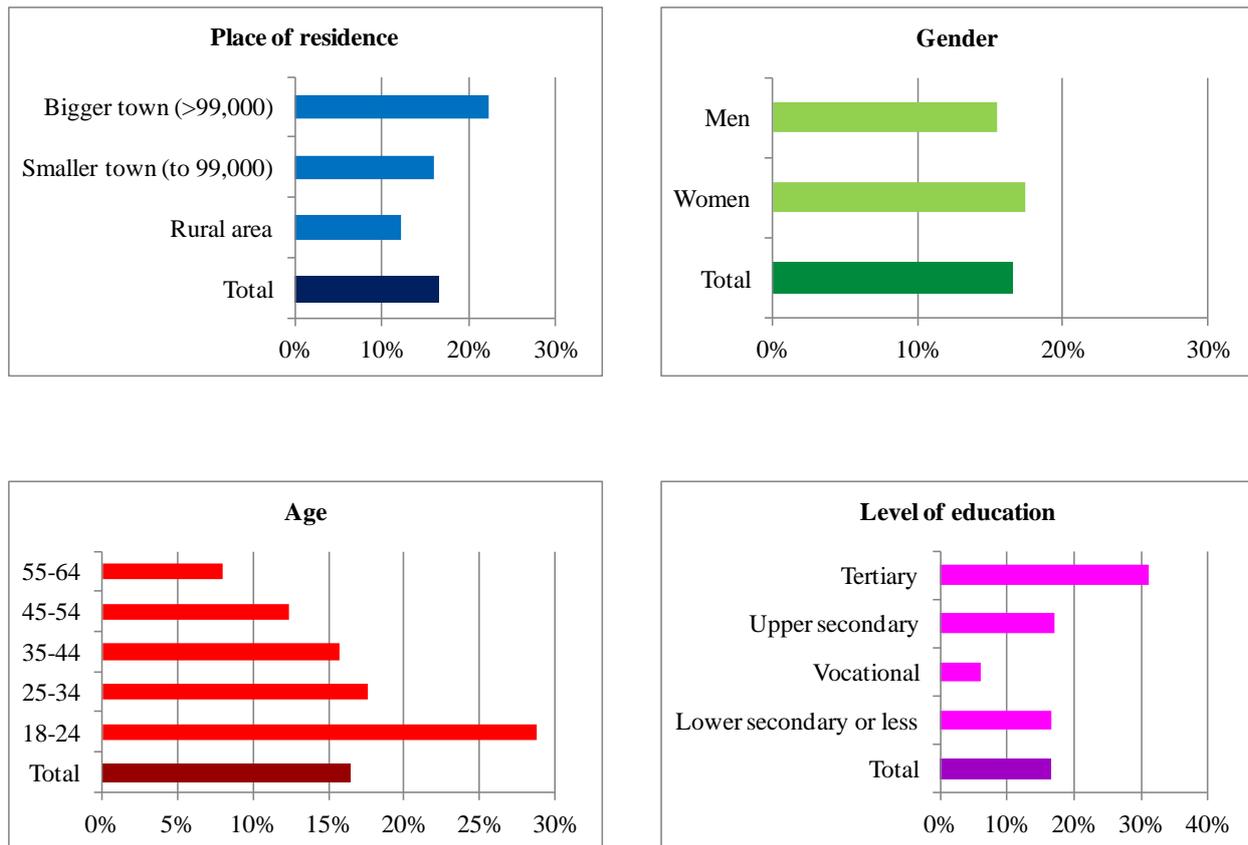


Figure 1. Declaration of undertaking self-learning activities in respect to various socio-demographic factors

Source: own computations on the basis of data from *Bilans Kapitału Ludzkiego - 2012*

Differences in the commitment to self-education are observed for all cross-sections. The most important findings are as follows:

- people from urban areas perform self-learning more frequently than those from rural terrains, moreover the frequency is higher in bigger towns and cities (i.e. more than 99 000 inhabitants) than in the smaller ones
- inhabitants of bigger towns and cities are more involved in such educational processes than inhabitants of smaller towns,
- women are more likely to self-learning activities than men, but the difference is not very large as it is equal to 2 percentage points,
- a strong relationship between the self-learning declaration and age of respondents is detected,
- the highest self-learning rate (28,8%) is observed for the youngest group, and generally, the older the person, the less chances for self-learning activities,

- there is a huge gap between those aged 18-24 and the others – the difference reaches up to 20 percentage points between the youngest and the oldest,
- self-education is extremely rare among persons with vocational level of education (6%) and quite popular among those with tertiary educational attainment (31,1%),
- generally, higher educational level inclines to undertaking self-learning activities, moreover there is a big gap between those with tertiary educational attainment and the others.

The conclusions drawn from the distributions of responses are confirmed by statistical measures calculated on the basis of two-way contingency tables constructed for the binary variable representing the participation (or its lack) and the categories of socio-demographic characteristics. Chi-square statistics with the corresponding p-values as well as contingency coefficients are shown in Table 1.

There is a strong evidence of the dependency between all considered socio-demographic factors and self-learning declaration (in the sense of chi-square statistic). Although the contingency coefficients are not very high they indicate that the strongest association is observed for level of education, then for age, then for place of residence and the weakest for gender.

Table 1. Association measures between self-learning declaration and various socio-demographic factors

Variable	Chi-square Statistic	p-value	Contingency coefficient
Place of residence	212,838	0,000	0,109
Gender	12,838	0,000	0,027
Age	551,831	0,000	0,174
Level of education	892,136	0,000	0,220

Source: own computations on the basis of data from *Bilans Kapitału Ludzkiego - 2012*

#### 4 Evaluation of co-occurrences of self-learning activities by association rules

The previous paragraph of this paper is focused on the socio-demographic determinants of undertaking self-learning activities in general, and this part of the study is oriented towards more detailed approach aiming at examining the elements of informal education and revealing particular patterns in learning methods. The purpose is to search for co-occurrences of self-learning activities with application of association rules analysis. Only simple association rules with one antecedent and one consequent are considered, i.e. the rules of the form *activity A*  $\Rightarrow$  *activity B*. Two commonly used quality measures are applied. Firstly, support measure reflecting the probability of the conjunction of activity A and activity B estimated as the number of respondents who indicated both activity A and

activity B divided by the total number of respondents. Secondly, confidence measure equal to conditional probability of B with given A calculated as the number of respondents who pointed out both activities A and B divided by the number of those who indicated activity A. Only association rules which satisfy at least 10% support and at least 10% confidence level are taken into consideration and presented in Table 2.

Table 2. Detected association rules and their characteristics

Antecedent		Consequent	Support (%)	Confidence (%)
Books, professional journals	=>	Computer programs, Internet	37,89	66,24
Computer programs, Internet	=>	Books, professional journals	37,89	59,63
Help from friends and colleagues	=>	Computer programs, Internet	11,02	56,38
Computer programs, Internet	=>	Help from friends and colleagues	11,02	17,35
Help from friends and colleagues	=>	Books, professional journals	10,81	55,32
Books, professional journals	=>	Help from friends and colleagues	10,81	18,91

Source: own computations on the basis of data from Bilans Kapitału Ludzkiego - 2012

The association rules results are often presented by visual means which allow to show the relationships in an attractive and clear way. The statistical software STATISTICA 10 which was used to carry out this analysis uses so-called network chart to illustrate schematically crucial rules as well as relative support measures. Such a visualization of the detected association rules is given in Figure 2.

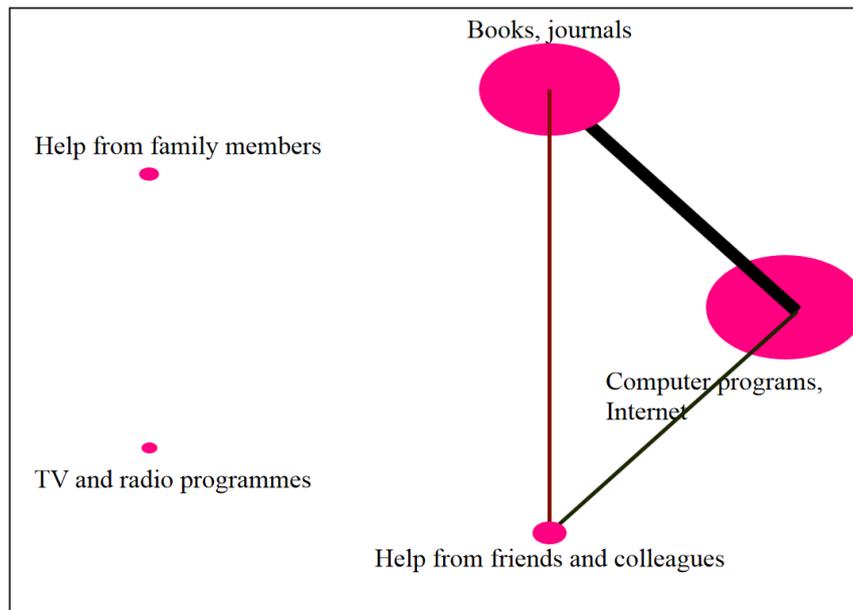


Figure 2. Visualization of the association rules among self-learning activities

Source: own computations on the basis of data from *Bilans Kapitału Ludzkiego - 2012*

The essential aspects of the visualization in Figure 2 are the size of the nodes reflecting the relative support of each element (i.e. activity) and the thickness of lines representing the relative support of two elements (activities).

The most common methods of self-learning declared by adult Poles are the use of computer programs and Internet as well as books, journals and other printed materials. Other ways of informal education are rather rare. It is worth underlying that activities supported by friend and colleagues are more frequent than those supported by family members. Only some respondents point out TV and radio programmes as a source of possible improvement of knowledge and skills. Activities undertaken in museums, exhibitions, galleries and science centers are hardly mentioned by adult Poles and therefore are not present on the network graph. The association rules which satisfy the established criteria concern three methods of self-learning: computer programs and Internet, books and journals, help from friends and colleagues. Interesting co-occurrences are found for each pair of the three aforementioned ways of informal learning what can be easily seen in the visualization chart where the lines form a triangle with vertices corresponding to these activities. The strongest evidence of interdependence is observed between computer programs and Internet usage and books and journals reading as the support measure is equal to 37,89% and the confidence measures are also very high (66,24% and 59,63%). It is worth noting that the confidence level is relatively high for the rules with the category "help from friends and colleagues" as antecedent. Respondents who benefit from the support given by friends and colleagues are also willing to use computer programs, Internet, books, journals and other printed materials.

## 5 Conclusions

Self-learning seems to be an advantageous way of improving skills and competences as it can be usually performed at the time convenient for the learner and it does not demand high costs. Despite this it is not a common method among adult Poles as only 16% of them declared in 2012 survey to have participated in informal learning activities in the last year. On the other hand, positive changes occur as the involvement rate increases. Self-education is not a homogeneous phenomenon when considered in respect with the socio-demographic characteristics. The educational level and age are the factors which contribute to the differentiation in the highest degree. Therefore, the groups which demand particular incentives are older persons with vocational level of education. Self-education processes in Poland are based essentially on two techniques - usage of modern ICT tools and more traditional way covering printed sources as books and journals. Moreover, these two methods have a high degree of interdependence. Certain importance is also observed for help of colleagues and friends. Computer programs and Internet are pointed out as the most frequent ways of improving knowledge and qualifications so the digital divide may be one of the crucial barriers to self-learning development.

### Acknowledgments

This research was financed by Narodowe Centrum Nauki (National Science Centre) in Poland under the project entitled "*Non-metric multivariate data analysis as a tool for study of adults situation in the context of demographic changes*". Project number: 2012/05/B/HS4/02499.

This paper uses data from the survey on human capital in Poland conducted in 2012 entitled *Bilans Kapitału Ludzkiego - 2012*.

### References

- Adults in formal education: Policies and practice in Europe*, 2011, Eurydice network.
- Bilans kapitału ludzkiego w Polsce. Badania ludności*, 2012:  
[http://bkl.parp.gov.pl/system/files/Downloads/20120917112349/BKL\\_kwest\\_ludno\\_3\\_edycja\\_final\\_.pdf?1347873882](http://bkl.parp.gov.pl/system/files/Downloads/20120917112349/BKL_kwest_ludno_3_edycja_final_.pdf?1347873882), retrieved 25-04-2014.
- Christensen, R., 1997, *Log-linear models and logistic regression*, Springer, New York.
- Cios, K. J., Pedrycz W., Swiniarski R.W & Kurgan L. A., 2007, *Data mining: a knowledge discovery approach*, Springer, New York.
- Classification of learning activities - Manual*, 2006, European Commission, Office for Official Publications of the European Communities, Luxembourg.
- Everitt, B. S., 1992, *The analysis of contingency tables*, Springer-Science+Business Media, B.V.
- Greenacre, M., 2010, *Correspondence analysis in practice*, CRC Press, Boca Raton.

- Grześkowiak A., 2013, *Statystyczna analiza aktywności edukacyjnej osób dorosłych w Polsce*, *Ekonometria* 2(40), pp. 22-35.
- Kryńska E. (ed.), 2013, *Instrumenty wspierające kształcenie ustawiczne pracowników. Regulacje i ich wykorzystanie w podmiotach gospodarczych*, Warszawa.
- Larose D. T., 2013, *Odkrywanie wiedzy z danych. Wprowadzenie do eksploracji danych*, Wydawnictwo Naukowe PWN, Warszawa.
- Lifelong Learning: the contribution of education systems in the Member States of the European Union*, 2000, Eurydice, the Information Network on Education in Europe, Brussels.
- Progress Towards the Common European Objectives in Education and Training. Indicators and benchmarks 2010/2011*, 2011, Publications Office of the European Union, Luxembourg.
- Szczucka, Turek & Worek, 2012, *Kształcenie przez całe życie*, PARP, Warszawa.