Evidence-Based Learning Approach in Evaluation of Information Literacy Education

Pavla Kovářová and Gabriela Šimková

Masaryk University, Faculty of Arts, Division of Information and Library Studies, Brno, Czech Republic
{kovarova,gsimkova}@phil.muni.cz

Abstract. The main aim of the paper is to present continuing research as a tool for more effective achievement of educational goals defined within information literacy (IL) education. Our primary resource is proven knowledge of the evidence-based learning approach, which we applied to information literacy education. Our article is designed as a series of case studies, each using different research methods inspired by Donald Kirkpatrick's Four-Level Evaluation Model. We focus on this particular model, because it corresponds with our aim to strengthen students' satisfaction and learning results. Each level of the model defines research objectives, achievable by application of different research methods. In our paper we present the practical application of selected research methods (implemented in IL lessons) to address each of the four evaluation levels. The described methods can be flexibly combined in order to obtain a complete picture of the real efficiency of investigated learning activity.

Keywords: Evidence-based learning, information literacy effectiveness, Kirkpatrick's Four-Level model, research.

1 Evidence-Based Learning

Various services labelled as "evidence-based" have been reported for some time and have been discussed with increasing frequency, among them: evidence-based medicine [1], evidence-based practice [2], evidence-based policy [3], but also evidence-based librarianship [4] and evidence-based teaching [5], evidence-based education [6] or evidence-based learning [7]. The origin of the approach is primarily identified with the first-mentioned field (medicine); including medical education [8], but the meaning has gradually shifted to the current delimitation: "the integration of professional wisdom with the best available empirical evidence in making decisions about how to deliver instruction", where professional wisdom is "the judgment individuals acquire through experience", empirical evidence is "scientifically-based research" and "empirical data on performance (is) used to compare, evaluate and monitor progress". [9] It follows from this that a combination of professional wisdom and research is necessary for effective education. If one of them is missing, the potential for effective education is not fully realised.

S. Kurbanoğlu et al. (Eds.): ECIL 2014, CCIS 492, pp. 560–569, 2014.

Even though the notion of evidence-based learning is relatively new as a term, its effects have been felt for much longer and also by those who have never heard of this concept. As pointed out by Davies: "The demands being made upon teachers and others who provide education call out for educational practice to be based on the best available evidence as well as the professional skills, experience, and competence of teachers." [6] Probably everyone who has lectured at several educational events has at least sometimes asked themselves whether they have managed to pass to their audience intended knowledge in such a manner that the audience is able to receive, retain and effectively use it. Asking this question is the first step toward evidenced-based practice. But if we also wish to get answers, we need to take the second step, which is to explore on our own initiative by engaging in research and reflection. This can bring positive effects not only for the teachers who will be able to improve their teaching, but also for the students. This was the central argument for providing nation-wide support for research in education in the USA by the No Child Left Behind Act [10].

The increasing emphasis on implementing student-centred research activities in the learning process points to the fact that the evidence of impact on the learning experience should be collected systematically. According to JISC such methodology "should display at least some of the following characteristics: be naturalistic (focusing on informal as well as formal learning), use more than one source of data, sample purposively (choosing learners who are characterised by behaviours or qualities of particular relevance), be mixed mode (e.g. diaries; observations; interviews; focus groups), employ semi-structured interview schedules and use open-ended methods (allowing unexpected issues to emerge)." [11]

The primary goal of our contribution is to stress the necessity of including cyclic or continual research methods in information literacy (IL) education. It seems that recent IL education may be characterized as specializing in a wide range of study areas. Thus, it is obvious that high-quality empirical data constitutes a precondition for developing effective student-centred lessons. Evidence-based learning in IL education leads to more effective achievement of educational goals.

As in other areas of research, the specific procedure to be applied in education research projects depends on the question that we ask. The approach in situations when we want to find out what the students already know and what we can safely build on without concerns about missing information is very different from a situation when we are interested in finding out whether students are able to put what they have learnt to practice. Each question has its merit, but provides only a partial picture of education. In order to get a more holistic view, various aspects of education evaluation are needed for a complete examination of the effects of education.

2 Kirkpatrick's Four-Level Model

Kirkpatrick's Evaluation Model is a specific approach to evidence-based learning. We focus on this model, because it corresponds with our aim to strength students' satisfaction and learning results. The practical experience with conducting research according to this approach is presented in detail.

Donald Kirkpatrick's model was first introduced in 1959 [18], primarily as a reaction to the demand for proving the effectiveness, value and benefit of education for business. It is presently one of the most widely used models for education evaluation. Its prevalence is also based on the fact that it reflects the current constructivist conception of instruction. Three main reasons for evaluation of education according to Donald Kirkpatrick are [11]:

1. To learn about the ways to improve future educational programmes
2. To decide whether to continue with an educational programme
3. To justify the very existence of an educational programme

The model is comprised of four hierarchically ordered levels revealing, one by one, the levels of effectiveness of the educational process. It thus enables an evaluator to try to answer the following questions: To what extent were the participants satisfied with the educational activity? To what extent did the participants obtain the expected knowledge and skills as a result of attending the educational activity? To what extent do the participants apply the knowledge and skills obtained to their everyday work? To what extent have the planned objectives of a development project and subsequent support activities been achieved? Donald Kirkpatrick's model is based on the assumption that "the desired result of the program is improved behaviour, with positive results to follow. If changes in behaviour result without measuring learning, there is no way to tell whether the change came from the training or from other sources. And if behaviour change is not evaluated, there is no way to tell whether the results came from the training program or from other sources." [12]

Kirkpatrick's model serves primarily as an education evaluation tool. Effective needs assessment is important also before the training [16], so that the content may be adjusted to a particular group (student-centred learning). It may turn out that failure of a lesson need not be caused by its quality in general, but rather by not meeting the participants' requirements. For example, it is evident that a top-quality university lecture on theoretical physics will not be effective for children starting their primary education. A needs analysis makes it possible to identify the characteristics of the students, and to tailor the lesson to their needs. Evaluation then makes it possible to subsequently observe any differences to determine whether they indicate a shift as a result of the education programme. This can be achieved by either a quasi-experimental or an experimental approach. The differences may be identified on all levels of the Four-Level model.

2.1 Immediate Reaction to Education

The first level of measurement tries to evaluate immediate student reactions to an educational activity (such as a seminar, a workshop, an e-learning module). Kirkpatrick conceived reaction as: "how participants feel about the various aspects of a training program." [17] He argued that a positive reaction is not a guarantee of effective learning but that a negative reaction guarantees refusal of (or non-participation in) the educational content.

Measuring the reaction is very easy to do and therefore it is also very often done, but not always in the right way, as when a survey does not meet quality research standards [17]. It is necessary to have a clear research goal, understandable questions, and quantifiable answers, to ensure the anonymity of participants, and it is desirable to include the possibility of respondents adding a comment. The result of level 1 research is a measure of participant satisfaction and motivation to learn about the topic. This level of evaluation is a subjective reaction, and this determines the research methods usually applied. They should be based on quantitative self-report by participants. The immediate reaction must be measured; and therefore it is necessary for respondents to fill in the sheet at the end of the lesson or very shortly after it.

Questions should reflect the defined research goals. For instance, Kirkpatrick includes questions pertinent to needs, ratio of lecture to discussion, 7 questions about the programme leader (e. g. keeping the session alive, using aids and illustrating points) and open questions on how to improve the lesson [17]. The IDEA evaluation is often used at universities for Kirkpatrick's level one evaluation [18].

We assessed students' satisfaction with the study environment, study content and the lecturer. In accordance with the research methods of Kirkpatrick's first level, we used short paper and pencil questionnaires (so called smile-sheets). We tried it both for children and adults. A simple version was prepared for 10–11 year olds. There are three emotion icons (emoticon) (smiling, neutral and sullen) and each child selects one of these emoticon when leaving the lesson to give feedback on how they liked the teacher, teaching methods and content by indicating which facial expression they would make if they should either come to the same lesson again or to evaluate it for a friend. Children expressed their reaction and added an oral comment on the teacher or classmates. This activity helped to convey the reaction of the children to the teacher; otherwise children would leave the lesson without expressing their reaction.

Another type of tool displaying immediate students' response to the educational activity just delivered were smile-sheets designed to be used in one-off information education lessons that took place at Masaryk University and were aimed at a target group comprised of students mainly aged between 20 and 25. The design of the questionnaire was based on an adjusted five-point Likert scale, which, instead of an evaluation reaching from "extremely satisfied" to "not at all satisfied", consisted of five emoticon indicating the level of satisfaction with a particular aspect. The three main aspects evaluated by this questionnaire were: content and organization of the seminar, the instructor and overall assessment of the lesson. A significant component was also an open question allowing participants to freely express their opinion on anything that they consider essential immediately after the lesson.

The purpose of the survey was to identify organizational and content problems which can be reduced to improve the lesson. This can lead to better students reactions and motivation to learn. But good results at this level are not enough – a lesson can be easy and enjoyable but with no learning effect. Therefore, it is necessary to continue to the next level of the model.

2.2 Knowledge Gained

The second of Kirkpatrick's levels explores the change in one or more areas of participants' knowledge, skills or attitudes due to an education activity [17]. This change is expressed by the quantity of knowledge transferred during a lesson, and therefore it should also be measured immediately after the lesson. On the other hand, the knowledge a participant is able to retain can be different. Therefore, the notions of immediate retention and delayed knowledge retention are sometimes employed [19].

On this level, Kirkpatrick [17] again recommends quantitative methods with statistical evaluation when knowledge is measured both prior to and after the lesson for a comparison. Further comparison should be performed with the use of a control group. The measuring should be as objective as possible. These planning and evaluation procedures require knowledge of measurement theory and statistical methods. Pre- and post-tests of various types (sometimes standardized) are applied. The aim of measuring at this level is to find out the amount of knowledge acquired by individual participants.

In order to map the progress of knowledge in the information literacy e-learning course, we assigned a pre-test and a post-test. The tests contained a series of 22 multiple-choice questions [only one answer]. In addition to a choice from three possible answers, the students were offered also the possibility of answering "I don't know", the inclusion of which helps to reduce distortion of test results through guessing. The students were urged to answer only when they knew the right answer; otherwise they were asked to select the "I don't know" option. The tests were fully anonymous. The aim of the tests was to map the initial knowledge of the group of students attending the course and subsequently to get an idea of the knowledge with which they left the semester course. The pre-test and post-test design often includes testing with a control group, ensuring a strong level of internal validity. "The principle behind this design is relatively simple, and involves randomly assigning subjects between two groups, a test group and a control group. Both groups are pre-tested, and both are post-tested." [13] The difference between the groups on the post-test, controlled by the difference on the pre-test will indicate whether the treatment group were more successful in the education process. However, in our case a control group was not used. The need for its employment is also mentioned by Naugle [18]: "The final guideline offered for evaluating training for resulting behavioural change was that a control group, not receiving the training, is to compare the difference between those who received instruction and those who did not. While we obviously cannot let a group of traditional students go a semester or a year without instruction to verify the benefit of instruction, we can utilize other forms of statistical controls." The tests examined the areas of information literacy that are taught on the course – reaching from information search through analysis and evaluation of information through to publishing. The results of the testing showed measurable progress in students' knowledge.

Another method was a didactic test which served the purpose of comparing the knowledge of both LIS students and librarians who took part in the lessons on information safety. The electronic test with a few questions about completed education dealing with the topic was not connected with a concrete lesson, but it was used to gain an overview of the knowledge of respondents concerning information

safety, which was a topic covered in seminars held during several previous terms. It was sent universally to all available librarians and LIS students, not only those who completed some kind of education dealing with the topic, and therefore those who had not complete such education formed a non-random control group. The results were statistically processed. This test confirmed different knowledge in the group that received the instruction and in the control group and identified weak topics which should be taught more extensively. Secondary findings concerned motivation and participants' characteristics connected with better results in the test.

Level 2 evaluation is often the end of the evaluation process [16], but it is still not complete. When someone has knowledge but is not able to use it, we cannot speak about effective learning. That leads us to the next level.

2.3 Long-Term Effects

The subject of the third level of Kirkpatrick's Model is to identify the long-term change in participants' behaviour (e.g. three to six months after the lesson). Not even a first-rate educational programme will ensure that the knowledge gained will have an impact on the results, without being intentionally, continuously and consistently supported and consolidated. Students cannot change their behaviour before they have a chance to do so. It is equally impossible to predict when such behaviour change will occur.

The methods used for third-level measurements usually have a qualitative character. Among other things, this is due to the fact that "it is difficult to define standards that can be used for measuring the application of learning and the question whether any changes in behaviour can be attributed to training (or some other factors) can always be asked" [16]. The methods for measuring the long-term effects usually include an interview to find out the views of students or teachers who work with the person who attended the course. These people can report any change in the behaviour of the trainee using methods such as 360-degree feedback. In addition self-evaluation can be used and observation methods such as mystery 'shopping' or 'calling' techniques.

Qualitative methodology, specifically a focus group aimed at identifying the ability and willingness of the students to utilize the acquired knowledge and skills effectively, was used to evaluate students' attitudes toward an information literacy e-learning course. Another method was the 360-degree feedback realized in the form of six interviews.

The above-mentioned focus group series was carried out as part of an e-learning course called Information Literacy Course. The aim of the series was to find out whether the students have consolidated their knowledge and skills and whether a change in their behaviour has occurred. Given the nature of the course, this element was evaluated in particular with regard to changes in their attitude to searching for scholarly information, its evaluation and sorting, as well as producing scholarly texts and their presentation. A detailed scenario examining all the aspects connected with the passage of students through the course was prepared for the focus groups. The focus groups took place six months after the end of the semester course. This was based on the assumption that it is a sufficient period of time for the students to absorb

the acquired knowledge and skills and to demonstrate using them in practice. There were three groups of students interviewed, each group consisting of 7–10 participants. The discussion was moderated by an experienced moderator whose role consisted primarily in generating and maintaining group dynamics, which is essential for active sharing of information by a focus group. The focus groups showed that the majority of the students who completed the course were satisfied and, most importantly, revealed that students act more cautiously when handling information on an everyday basis, including information search and evaluation practices. Furthermore, students view themselves as more cautious when selecting resources for their term papers and final theses. A change in behaviour was thus apparent in several areas taught on the course.

Six interviews were conducted to obtain 360-degree feedback about an internet safety lesson provided by a library. The lesson was prepared for children in cooperation with their school. In the interviews, a teaching librarian, the director of the library, the deputy director of the school, a teacher, a pupil and her mother were asked about their evaluation of the lesson. The interviews were held 6 weeks to five months after the lesson. Questions were directed at the behaviour of pupils a few days after the lesson and also at the actual time. Interviewers agreed that the pupils spoke a lot about the topic immediately after the lesson. Discussion of the topic then fast decreased and the pupils were not able to repeat what was the concrete content of the lesson, but in problematic situations they remembered parts of it and were able to identify risk behaviour and realize possible consequences of their behaviour.

If data are collected conscientiously and reflect the reality under examination well, they can be: “used to assess students’ needs and encourage educators to meet those needs. The data can also be used to evaluate teachers’ efforts to meet those needs.” [18]

2.4 Results

The fourth level shows the tangible results of a programme and is accepted mainly in the commercial sphere because it is focused on the return on investment in education (in general). This is the only level that has not been performed independently and is therefore presented mainly theoretically.

The fourth level evaluates the overall effectiveness of instruction, the outcome in commercial sphere or the benefit for an institution brought about by a student's improved performance. This process is both time-consuming and financially demanding, but it is useful from a comprehensive point of view. The objective is to show tangible results of a programme. These results may include such factors as quality improvement, cost reduction or return on investment in education. The theory of return on investment in education was developed as an extension of Kirkpatrick's model by Phillips [15]. It is “a comprehensive approach for measuring the effectiveness of training that begins with planning the project, and then goes on [sic] techniques for collecting data and its analysis and finally, ends with a final report. In compares [sic] with the conventional financial ROI, as the ratio of earnings and investments, ROI for learning and development takes into account earnings as net

benefits from education programs (monetary difference between benefits and costs of the program) and investments as the real costs of the program.“ [14]

The process of evaluating an educational activity at the third and fourth level of Kirkpatrick's Model is best able to explore the educational effects of information education. This concerns not only students' satisfaction with the course or the knowledge and skills they acquire, but also the long-term changes noticeable both at an individual level and on a society wide level. Educational effects can be thus understood as long-term consequences and effects produced as products of educational processes. They affect the lives of individuals as well as the entire society in its economic, political, cultural and other characteristics – such as the impact of education on professional fulfilment and income of individuals, their cultural and political orientation, media preferences, leisure activities, consumer behaviour, labour productivity, and experience of unemployment.

We can also look at the relationship between the third and the fourth level – if the participants change their behaviour and consistently apply the acquired knowledge and skills in practice, productivity of individuals as well as of the whole enterprise increases.

Interviews in the 360-degree feedback mentioned above mapped also the opinion of interviewed managers and a parent (hierarchically higher levels). We consider that this data contributes to research on the fourth level. In order to cover this level, questions concentrating on actual effects of the change on the environment were included. However, in the interviews, these questions were added as a supplement to the semi-structured interview scenario only for half of the respondents, thus providing only a limited measurement of the fourth level. Since it was a lesson for children, it was difficult to identify the effects on the environment. Despite these limitations, some effects were identified. These were described by one interviewee as cultivation of children’s online environment experience, including active dealing with problems and possible solutions to them that the child is familiar with, instead of ignoring them. There was also a secondary effect of a moderate increase in reputation of the library in the eyes of all interviewees as a result of helping to address this social problem.

The advantages of the fourth level of measurement are reflected in determining the outcomes of training and its connections with educational goals [14].

3 Conclusion

The evidence-based approach in the field of IL education brings constant feedback about the ongoing education to the lecturer, supplies valuable data revealing the effectiveness of individual forms of IL education and helps to predict the course of further education activities. Data processed to a high standard can be used also as a compelling argument to substantiate the effectiveness of the costs expended or to support a potential application for granting additional resources for development of information education activities. Our contribution stresses the positives of application of Kirkpatrick’s model to designing effective IL education. To implement the model, various research activities were conducted. Measuring at each level brings partial results, which when combined assist in developing educational efficiency, but in the

case of getting the results together, new emergent evidence appears. The core goal of our paper is to demonstrate, that the Kirkpatrick's model application to IL education is not only a theoretical concept, but a flexible practical attitude which can bring many positives to both educator and students.

The first level tried to evaluate immediate student reactions to an educational activity (environment, content and the lecturer). Short paper questionnaires (smile-sheets) showed that some aspects of a lesson organization could be rethought (e.g. classroom equipment or air conditioning). The second level explored the change in knowledge and skills, usining a pre- and a post-test and a didactic test. The results showed that themes requiring deeper attention were identified. Qualitative methodology, specifically focus groups and a 360-degree feedback based on six interviews, were used within the third level to identify the long-term change in participants' behaviour. These aspects, which were identified as positive and long-lasting for students' learning, were enriched (e.g. effective multimedia materials, such as videos and audio podcasts, were accomplished by webinars). The fourth level is focused especially on measuring of economical, including financial benefits including reduced costs, higher quality, and lower rates of absenteeism [17]. The transfer of children's knowledge to a productive population (parents) was verified.

Smith [10] emphasizes one of the problems of evidence-based learning is the absence of a platform for researching and sharing of good practice in lessons, as well as for benchmarking and establishing contacts for assistance in introducing research into education.

Although most teaching librarians are not research experts, some are and the field of IL is not so vast and, most importantly, Europe is not so diverse that segmentation is necessary. On the contrary, it would be advisable to join forces and share good practice. Perhaps collaborators for creating an online platform available as and when needed to everyone interested will appear, or this role will be fulfilled by conferences such as ECIL. In any case, it makes sense to build such a database, and to strive for research-based resources with empirical evidence of their effectiveness. This could lead to sharing of good practice in order to achieve what we are all concerned with – improving information literacy in society.

References

1. Evidence-based Medicine. JAMA 17, 2420 (1992)
2. Olson, E.A.: Evidence-Based Practice: A New Approach to Teaching the Integration of Research and Practice in Gerontology. Educational Gerontology 22(6), 523–537 (1996)
3. Leicester, G.: The Seven Enemies of Evidence-Based Policy. Public Money & Management 1, 5–8 (1999)
4. Eldredge, J.: Evidence-Based Librarianship: Searching for the Needed EBL Evidence. Medical Reference Services Quarterly 3, 1–18 (2000)
5. Mitchell, D.: What Really Works in Special and Inclusive Education: Using Evidence-based Teaching Strategies. D. Mitchell, London (2008)
6. Davies, P.: What is Evidence-based Education? British Journal of Educational Studies 2, 108–121 (1999)

7. Cranney, J.: Toward Psychological Literacy: A Snapshot of Evidence-based Learning and Teaching. Australian Journal of Psychology 1, 1–4 (2013)
8. Davies, P.: Approaches to Evidence-based Teaching. Medical Teacher 1, 14–21 (2000)
9. Whitehurst, G.J.: Evidence-based Education (EBE), http://ies.ed.gov/director/pdf/2002_10.pdf
10. Smith, A.: Scientifically Based Research and Evidence-Based Education: A Federal Policy Context. Research & Practice for Persons with Severe Disabilities 3, 126–132 (2003)
11. Kirkpatrick, D.: The Four Levels of Evaluation: Measurement and Evaluation. American Society for Training & Development Press, Alexandria (2007)
12. Kirkpatrick, D.: Seven Keys to Unlock the Four Levels of Evaluation. Performance Improvement 45, 5–8 (2006)
13. Explorable: Pretest-Posttest Designs, https://explorable.com/pretest-posttest-designs
14. Rahimić, Z., Vuk, S.: Evaluating the Effects of Employees Education in B&H Companies. In: International Conference of the Faculty of Economics Sarajevo (ICES), Sarajevo, pp. 1044–1057 (2012)
15. Phillips, J.J., Phillips, P.P.: The ROI Quiz: The Myths & Mysteries of ROI (2006), http://www.roiinstitute.net/m/uploads/articles/pdf/2007/06/19/MythsandMysteries.pdf
16. Kirkpatrick, J.: The Hidden Power of Kirkpatrick's Four Levels. T+D 61(8), 34–37 (2007)
17. Kirkpatrick, D.: Revisiting Kirkpatrick's Four-level Model. Training and Development 1, 54–59 (1996)
18. Naugle, K.A., Naugle, L.B., Naugle, R.J.: Kirkpatrick's Evaluation Model as a Means of Evaluating Teacher Performance. Education 1, 135–144 (2000)
19. Haupt, G., Blignaut, S.: Uncovering Learning Outcomes: Explicating Obscurity in Learning of Aesthetics in Design and Technology Education. Int. J. Technol. Des. Educ. 18, 361–374 (2008)