

An Overview of Research Methods of Gain

K. Kolář

Charles University in Prague, Faculty of Mathematics and Physics, The Czech Republic.

Abstract.

The paper describes some suitable methods of educational research for determination of the gain of an activity, i.e., what students/participants of some course/activity learned during that. An accurate approach to measure gain is an usage of pre- and post-test, interview or questionnaire — the article briefly summarizes these methods. There are often used concept tests/inventories which should reveal the misconceptions of tested students. Further, we introduce our plans on the translation of the Calculus Concept Inventory into Czech and the intention to use the described methods on study of gain in *Science Centres*.

Introduction

There are many possibilities for the public to meet with some science outreach activities (activities popularizing science). Very popular are *Science Centres*, which are in essence very interactive science museums. Primary and secondary school pupils can also participate in many competitions which should stimulate their interest in science. In central Europe, there is a “local speciality”¹ called *Correspondence seminars*. There are also one-day-events like the One Day with Physics (*Jeden den s fyzikou*), the *Masterclasses* and many others. The year 2015 is declared as The Year of Industry and Technical Education² in the Czech Republic and society is encouraged in development and participation in such events.

All activities mentioned above should contribute to teach participants something new and changing the attitude of participants and motivate them — for example for choosing or even changing the field of their next study. Very interesting topic to study is “efficiency” of these activities. Their efficiency has three parts — gain (in the meaning described further), change of attitude of participants and economical aspect. All three aspects are important but we'll consider only the first one in this article.

This article is a literature review focused on methods which can be used for measurement of what the participants actually learned (called often gain), especially in the field of physics. Research in the science education is concentrated on the schools and universities but not so much to out-of-school activities. Therefore the methods mentioned further originate from researches done in schools or in psychology.

There are also more methods how to estimate what did the participants learn (more for instance in [Cumming, 2012]). We can investigate for example *effect size*. It can be used, for example, *Cohen's d* or *eta-squared*. But those numbers depend on the amount of participants — the effect size grows with number of participant and this article is mainly about gain.

Some of described methods will be selected for the next study concerning the gain of Science Centres, FYKOS' camps³ and Masterclasses.

¹As far as author knows these activities are typical for The Czech Republic, Slovakia and Hungary.

²Web page: <http://www.rokprumyslu.eu/>.

³FYKOS is abbreviation from *Fyzikální korespondenční seminář* — this name could be translated as *Physics Correspondence Seminar*. It is an activity for gifted high school students in Physics. The web pages of this activity are <http://fykos.org>.

The main approaches and terms

Definition of gain

The word *gain* has many slightly different meanings in English, so we define it for our purposes here. For standardized test, we can use definitions from [Hake, 1998]

- *Average gain* $\langle G \rangle$ for a course (analogically for an activity) is defined as

$$\langle G \rangle = \langle S_{\text{fin}} \rangle - \langle S_{\text{in}} \rangle,$$

where $\langle S_{\text{in}} \rangle$ is average initial score (from pre-test) and $\langle S_{\text{fin}} \rangle$ is final class (group) average (from post-test).

- *Average normalized gain* $\langle g \rangle$ for an activity is ratio of the actual average gain $\langle G \rangle$ to the maximum possible average gain $\langle G \rangle_{\text{max}} = S_{\text{max}} - \langle S_{\text{in}} \rangle$, where S_{max} is the maximum possible score. We can write

$$\langle g \rangle = \frac{\langle G \rangle}{\langle G \rangle_{\text{max}}} = \frac{\langle S_{\text{fin}} \rangle - \langle S_{\text{in}} \rangle}{S_{\text{max}} - \langle S_{\text{in}} \rangle}.$$

This definition is sometimes called Hake's average normalized gain.

We can modify these definitions for individual ones this way

- *Individual gain* G_i for a student (or a participant) is

$$G_i = S_{i,\text{fin}} - S_{i,\text{in}},$$

where $S_{i,\text{in}}$ is participant's pre-score and $S_{i,\text{fin}}$ is participant's post-score.

- *Individual normalized gain* g_i for a participant is

$$g_i = \frac{S_{i,\text{fin}} - S_{i,\text{in}}}{S_{\text{max}} - S_{i,\text{in}}}.$$

Individual gain and individual normalized gain can be used only for information for the specific participant or for qualitative research approach. For quantitative approach are used only the average gain and the average normalized gain.

Pre-test has to be prepared in the way that no pretested subject would receive S_{max} , unless he/she has absolved similar activity. In the case that $S_{i,\text{in}} = S_{\text{max}}$, the normalized gain could not be even defined. The post-test should be also stratifying in accordance to what the participant really learned. The tests used as pre-test and post-test has to have high validity and reliability so that the gains have the intended meaning and are not only some random numbers (due to high error of measurement).

For usage in this article, we generalize the meaning of word *gain*. *Gain* of activity is what participants of some activity learned thanks to take part in this activity during that activity. This enables us to use this word also in the context of the qualitative research methods. In some articles, the term *added value* could be used instead of the term gain.

Description of literature search

Research methods were searched in literature by several means described in the list below.

- There were at first searched some general oriented publications like [Fraser et al., 2011], [Bell, 2010], [Gavora, 2010] and [Ferjenčík, 2000] for making a basic overview of methods often used in the education and the psychology research.

- Several searches were done on the *Web of Science*TM ⁴, one of them is described here. Purpose of this search was to find what methods were really applied on research of physics activities in recent years. Also it was a way to find some new and not yet cited articles. More searches were conducted, but we present here only this one. It's parameters were: *topic* — gain, physics, education method; *years of publication* — 2005–2014. In sum, 48 results were found and all abstracts of them were analysed. Some of them were not relevant for this topic, some of them were using the same methods and some of them are cited further.
- Other searches were also conducted on the *Web of Science*TM and the *Google Scholar* concerning specific activities such like “Science Centres” and “Masterclasses.”
- Some contributions of previous years of Week of Doctoral Students in the branch f12 were searched. Also some additional resources were added on basis of consultation with colleagues from Department of Physics Education and communication with Czech Science Centres itself.

Results

From the various range of research methods those were selected that can us give information about (average or individual) gain.

Quantitative vs. qualitative research

As already mentioned, the gain is well defined for quantitative methods. We can use for it definitions for average gain and average normalized gain. If we have sufficiently large and representative sample, we can generalise and predict for activities under similar conditions. But if we decide to use quantitative approach we loose particular information about individuals, we cannot study context so deep as it could be studied by qualitative methods. So for our purposes it seems to be the best to choose mixed research, combining the qualitative and quantitative methods of research.

Observation

Observation is in its nature more qualitative method but if we use coding it could be more quantitative. It's not a good method for obtaining the average normalized gain because the determination, whether the participants really understood the concepts which were described to them solely from the observation, is complicated.

But this method shouldn't be completely discarded because it can be a useful complement to other methods. Observation is for example used in Science Centres, as author was informed by one of the employees of on of the Czech Science Centres, for determining whether concrete exhibit is popular and the visitors are working with the exhibit correctly. Science Centres often use some kind of an automatic evaluation of number of visitors visiting some exhibit each day and average time spent with the exhibit. They than evaluate it and change the composition of exhibits, modify them or exclude some exhibit from the exhibition.

Interview

Many kinds of interview do exist. In psychology, even technique using *open associations* can be called an interview as states in [Ferjenčík, 2000]. That's obviously the most qualitative interview which could be made. But more common is to divide interview in these next kinds (from the most qualitative to the most quantitative one) [Dawson, 2007]

⁴<http://apps.webofknowledge.com>

- *Unstructured interview* or *in-depth interview* — there are no prepared questions for such interview, only a general area of interest, the interviewer has to be as adaptable and open as possible has to let talk the interviewee freely.
- *Semi-structured interview* — there are prepared several general questions or areas, interviewer gives the interviewee some liberty but in the case that the interviewee is turning away from the topic, interviewer asks question about topic of interest and returns to the prepared questions or area.
- *Structured interview* — all the questions are prepared. The questions can be either open-ended or even close-ended. Researcher is just asking for the specific prepared answers and is leading the interview.

To be an interview is considered often just face to face conversation among interviewer and interviewee but [Ferjenčík, 2000] points out that an interview could be also other kinds of communication between researcher and surveyed participant — for example questionnaires, tests.

If we want really to study gain, we have to make an interview before and after activity. If only interview afterwards is possible, we can at least qualitatively describe at least what the interviewed participant thinks he/she learned (but not actually what he/she precisely learned).

For example, the research of [Falk, Storksdieck, 2005] uses the standardized interviews and the observation of visitors of Science Centre.

Questionnaire and tests

Questionnaires could be regarded as kind of standardized or structured interview [Ferjenčík, 2000]. However the researcher automatically loses an information about the participant which could be recorded in interview, like the facial expression of interviewee.

Many kinds of questions can appear in a questionnaire and the closeness of these questions determines the degree of quantitiveness of the questionnaire.

There can be many kind of tests. The “best ones” are the standardized didactic tests. But there are numerous test made by teachers with the the objective to mark their students. These tests can be used for studying gain if they are validated and if they have sufficiently high reliability.

Concept inventories

Concept inventories or concept test are standardized test specialized in testing conceptual understanding of some topic. Concept test arise from constructivism. The theory states that the students bring to education their prior knowledge and experiences [Skalková, 1999]. These are usually called preconceptions. In the case that the preconceptions are not in accordance with the scientific theories, they are called alternative preconceptions or misconceptions.

List of some existing concept test can be found on [NC State University, 2015] (in English) or on [Projekt KEGA, 2015] (in Slovak).

Application of various concept inventories can be found in numerous articles, for example [Thacker at al., 2014] inform us about simultaneous usage of the Force Concept Inventory, the Brief Electricity and Magnetism Assessment, the Mechanics Baseline Test and the Conceptual Survey of Electricity and magnetism for several years.

Concept test are often used when the traditional teaching methods (T; like “talk and chalk”) and the interactive engagement methods (IE; like peer instruction [Crouch, Mazur, 2001]) are compared. Most of the articles emphasize that IE methods of instruction have greater gain than T. But article [Ding et al., 2008] warns that the results of an application of a concept test depend on the conditions used and article of [Franklin et al., 2014] further warns that the

difference between the T and IE is maybe more the duration of learned information. It claims that IE education has more long-term effect.

From the point of view of the informal activities, these test are often too long to be used there (more than 20 minutes) and are not adapted for the use of the pre-test and post-test the same day.

The Force Concept Inventory. The Force Concept Inventory is one of the most known concept inventories. It can also be one of the most used ones and became the concept test towards which are other concept test judged. It's topic is conceptual understanding of Newton's mechanics.

Evidence that with this concept inventory is still worked, is for example the article of Dividing the Force Concept Inventory into two equivalent half-length tests [Han *et al.*, 2015].

The Thermal Concept Evaluation. The Thermal Concept Evaluation is one of the concept inventories and is focused on detection of misconceptions or false preconceptions in thermodynamics. It was translated and piloted by our faculty in 2013 [Káčovský, 2013].

The Calculus Concept Inventory. The Calculus Concept Inventory (CCI) tests the conceptual understanding of basics of calculus — especially the derivatives [Epstein, 2007], [Epstein, 2013]. We have obtained the original English version an permission for translation into the Czech from the author.

Future plans

Our team would like to use standardized interview and questionnaires in the Science Centres. At first, we are going to concentrate more on the scientific program which are Science Centres preparing for schools. The ideal is to make pre- and post-visit interviews with representative sample of regular visitors of one of the Science Centres in the Czech Republic. But the used questions should be as short as possible, due the relatively low motivation of visitors for long answering in such surveys. The questions should be similar to the ones in concept inventories because we would like to determine whether the exhibition leads to the conceptual understanding.

The CCI is going to be used on FYKOS' camps. The reason for that is because the derivatives are one of the basic math skills which is need to be mastered for mastering of college level of physics which is often taught there. Later we'd like to apply it also in calculus courses at our faculty.

Conclusion

This paper summarized some useful methods for the gain investigation. Most used and probably most valid and reliable are the concept tests. These are important also because of their approach to the revealing conceptual (mis)understanding of the chosen topic. But the concept inventories has to be correctly used and the conditions of testing has to be described well. As a result, we are going to translate the Calculus Concept Inventory and use it on FYKOS' camps for high school students and calculus courses on our faculty.

Acknowledgments. The author thanks Jiří Dolejší, Vojtěch Žák and Hana Šustková for many critical remarks about this article. The study was supported by the Charles University in Prague, project GA UK No 188515.

References

- Bell, J.: *Doing Your Research Project — A guide for first-time researchers in education, health and social science*, 5th edition, Glasgow, ISBN-13 978-033523582-7, 2010.
- Crouch, C. H., Mazur, E.: Peer Instruction: Ten years of experience and results, *Am. J. Phys.*, **69**, 970, 2001.

- Cumming, G.: *Understanding the new statistics: effect sizes, confidence intervals, and meta-analysis*, Routledge, New York, ISBN 978-020380700-2, 2012.
- Dawson, C.: *A Practical Guide to Research Methods; A user-friendly manual for mastering research techniques and projects*, 3rd edition, Glasgow, ISBN 978-184803223-1, 2007.
- Ding, L., Reay, N. W., Lee A., Bao, L.: Effects of testing conditions on conceptual survey results, *Phys. Rev. Sp. Top. — Phys. Ed. Res.*, **4**, 010112, 2008.
- Epstein, J.: Development and validation of the Calculus Concept Inventory, in *proceedings of the Ninth international conference on mathematics education in a global community*, 165–170, Charlotte, NC, 2007.
- Epstein, J.: The calculus concept inventory-measurement of the effect of teaching methodology in mathematics, *Notices of the American Mathematical Society* 60.8, 1018–1027, 2013.
- Falk, J., Storksdieck, M.: Using the contextual model of learning to understand visitor learning from a science center exhibition, *Sci. Ed.*, **89**, 744–778, 2005.
- Ferjenčík, J.: *Úvod do metodologie psychologického výzkumu*, Portál s.r.o., Praha, ISBN 80-7178-367-6, 2000.
- Franklin, S. V., Sayre, E. C., Jessica, W. C.: Traditionally taught students learn; actively engaged students remember, *Am. J. Phys.*, **82**, 798, 2014.
- Fraser, B. J., Kenneth, T., Campbell J. McR. (editors): *Second international handbook of science education*. Vol. 24. Springer Science & Business Media, 2011.
- Gavora, P.: *Úvod do pedagogického výzkumu*, Paido, Brno, ISBN 978-80-7315-185-0, 2010.
- Hake, R. R.: Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses, *Am. J. Phys.*, **66**, 64, 1998.
- Han, J. et al.: Dividing the Force Concept Inventory into two equivalent half-length tests, *Phys. Rev. Sp. Top. — Phys. Ed. Res.*, **11**, 010112, 2015.
- Kácovský, P.: Students' Alternative Conceptions in Thermodynamics, in *WDS'14 Proceedings of Contributed Papers — Physics (eds. J. Safrankova and J. Pavlu)*, Prague, Matfyzpress, pp. 100–103, 2013.
- North Carolina State University Department of Physics: *Assessment Instrument Information Page* [online], available at <http://www.ncsu.edu/per/TestInfo.html> cited on 29th May 2015.
- Projekt KEGA: Databáza konceptuálnych a postojových testov, *Interaktívne prednáškové demonštrácie a Peer Instruction — nové interaktívne metódy zvyšujúce efektívnosť prednáškového formy vzdelávania* [online], available at <http://physedu.science.upjs.sk/metody/testy.html>, cited on 29th May 2015.
- Skalková, J.: *Obecná didaktika*, ISV nakladatelství, Praha, ISBN 80-85866-33-1, 1999.
- Thacker B., Dulli H., Pattillo D., and West K.: Lessons from a large-scale assessment: Results from conceptual inventories, *Phys. Rev. ST Phys. Educ. Res.* **10**, 020104, 2014.