EVALUATING FINANCIAL EDUCATION PROGRAMMES: SURVEY, EVIDENCE, POLICY INSTRUMENTS AND GUIDANCE

This document is submitted to the G20 Leaders' Summit on 5-6 September 2013. It addresses the G20 call to develop practical tools for the evaluation of financial education programmes. It was welcomed by G20 Finance Ministers and Central Bank Governors in July 2013.

For further information, please contact Mr. André Laboul, Head of the Financial Affairs Division, OECD; Chair of the INFE [tel: +33 1 45 24 91 27; E-mail andre.laboul@oecd.org] or Ms. Flore-Anne Messy, Senior Policy Expert, Financial Affairs Division, OECD and INFE Executive Secretary [tel: +33 1 45 24 96 56; E-mail: flore-anne.messy@oecd.org]

September 2013

FOREWORD

In June 2012 G20 Leaders recognised the importance of financial education and endorsed the High-level Principles on National Strategies for Financial Education. In the same occasion G20 Leaders also called on the OECD/INFE and the World Bank in cooperation with the GPFI to deliver further tools to promote financial education.

Understanding whether financial education works, how it works and what are the most appropriate methods for evaluating financial education programmes are key components of a successful national strategy for financial education. Good programme evaluation allows to show whether objectives are being met, to identify elements that can be scaled-up or replicated. It also allows policy makers to test different approaches to see which are the most cost efficient, and to assess whether different methodologies have differential impact on various population subgroups.

Under the support of the Russia/WB/OECD Trust Fund for Financial Literacy and Education, the OECD/International Network on Financial Education (INFE) has developed significant work in relation to evaluating financial education programmes, including the collection of countries' experiences, challenges, and lessons learnt in evaluating their financial education programmes. Based on these evidence and lessons, the INFE has then developed policy instruments on the evaluation of financial education programmes, including the INFE High-level Principles on Evaluation of Financial Education Programmes (2011) and two non-technical OECD/INFE Guides to Evaluation (2011).

This book collects country evidence and lessons learnt in evaluating financial education programmes, suggests an overall framework to guide policy makers and financial educators when designing an evaluation study, and contains the High-level Principles on Evaluation of Financial Education Programmes. Two non-technical OECD/INFE Guides to Evaluation are also included in the Annex. This publication complements the Toolkit for the Evaluation of Financial Capability Programs in Low- and Middle-income Countries developed by the World Bank under the Russia/WB/OECD Trust Fund for Financial Literacy and Education.

The content of this book has been developed by the OECD/INFE and approved by the OECD bodies responsible for financial education. It was further circulated and discussed during the conference "Empowering Consumers of Financial Products and Services through Improved Financial Literacy and Effective Consumer Protection Systems" held on June 13-14, 2013 in Moscow and organised under Russia's G20 Presidency with the support of the World Bank and the OECD. G20 Finance Ministers and Central Bank Governors welcomed the work of the OECD/INFE and the World Bank on the development of practical tools for financial literacy measurement and the evaluation of financial education programmes at their meetings in February and July 2013. The content of this book also complements the World Bank "Toolkit for the Evaluation of Financial Capability Programs in Low- and Middle-income Countries".

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This book draws together expert opinions and recommendations on the evaluation of financial education programmes. It provides evidence of countries' experiences, challenges, and lessons learnt in evaluating their financial education programmes, and suggests an overall framework to guide policy makers and financial educators when designing an evaluation study. Finally, the book contains the High-Level Principles for the Evaluation of Financial Education Programmes that were developed by the OECD/INFE based on the above evidence and analysis.

The content of this book was developed by the OECD/INFE and approved by the OECD bodies responsible for financial education. It was further circulated and discussed during the conference "Empowering Consumers of Financial Products and Services through Improved Financial Literacy and Effective Consumer Protection Systems" held on June 13-14, 2013 in Moscow and organised under Russia's G20 Presidency with the support of the World Bank and the OECD.

Chapter 1 focuses on the actual processes that have been used to evaluate existing financial education provision in various countries. It presents the results of a comprehensive stock take of monitoring and evaluation activities undertaken in 2009 in all countries within the OECD International Network on Financial Education (INFE), collecting responses from 46 authorities in 29 countries.

The chapter begins with a brief introduction to explain why evaluation is important, before describing the wide variety of financial education programmes that have been evaluated and the range of methods used in evaluation. It then highlights the challenges and lessons learnt mentioned by the responding authorities in their own attempts to evaluate, including the need to take independent evaluators; to undertake continuous and regular monitoring and evaluation of the programme; to plan the evaluation alongside the design of the initiative; and to identify appropriate evaluation and control participants.

Overall, the stock take revealed that many authorities were undertaking monitoring and evaluation on a regular basis, but that each of them had made its own decision about how to evaluate, which methods to use, who to ask for guidance, and what aspects of the education process to focus on. This points to the benefits of a common framework allowing countries flexibility to evaluate individual schemes while at the same time using well tested methods.

Chapter 2 suggests an overall framework to guide policy makers and financial educators when designing an evaluation exercise. In particular, it highlights the main benefits of programme evaluation, along with the challenges faced by evaluation designers. For example, resources can be used more effectively and the overall impact of financial education can be increased by using evaluation data to make decisions about which programmes should be improved, which should be continued and whether any should be terminated.

The chapter discusses the challenges faced by evaluators, and suggests a framework that allows significant flexibility whilst still providing standardisation and comparability. The chapter also recognises that there are potential pitfalls when such a framework is used, including the need to have clearly defined objectives and to define the length of evaluation time.

The chapter ends with recommendations and conclusions, stressing that:

- Evaluation is important and should be designed alongside the financial education programme;

- Employing quasi-experimental methods (ones that mimic scientific experiments in medicine, or agriculture, for example) will give the evaluation added credibility and help to assess causality;

- Employing professional evaluators, using evaluation toolkits/frameworks and subjecting evaluation design and reporting to peer-review can all ensure that evaluations are robust and comparable.

Chapter 3 contains the High-Level Principles for the Evaluation of Financial Education Programmes that were developed by the OECD/INFE based on the above evidence and analysis. These High-level Principles were developed with financial support from the Russian/World Bank/OECD Trust Fund on Financial Literacy, and approved by the OECD Committee on Financial Markets (CMF) and the OECD Insurance and Private Pensions Committee (IPPC) in 2011.

Two non-technical OECD/INFE Guides to Evaluation are included in the Annex. These short guides complement the High-Level Principles and offer practical tools to plan and implement the evaluation of financial education programmes, as well as to disseminate its results.

Chapter 1

ANALYTICAL REPORT ON EVALUTION OF FINANCIAL EDUCATION PROGRAMMES

This chapter focuses on the processes that have been used to evaluate existing financial education provision in various countries. It is the result of a comprehensive stock take of monitoring and evaluation activities in all countries within the OECD International Network on Financial Education (INFE), undertaken in 2009. The chapter begins with a brief introduction to explain why evaluation is important, before describing the wide variety of financial education programmes that have been evaluated and the range of methods used in evaluation. It then draws conclusions from the stock take of current evaluation practices.

Introduction

This chapter focuses on the actual processes that have been used to evaluate existing financial education provision in various countries. It is the result of a comprehensive stock take of monitoring and evaluation activities in all countries within the OECD International Network on Financial Education (INFE), undertaken in 2009 on behalf of the INFE Expert Subgroup on Programme Evaluation.

The chapter begins with a brief introduction to explain why evaluation is important, before describing the wide variety of financial education programmes that have been evaluated and the range of methods used in evaluation. It then draws conclusions from the stock take of current evaluation practices.

The complementary roles of monitoring and evaluation

When we talk about the evaluation of a financial education programme, we are really discussing two distinct, though not entirely separate, processes. The first is the process of monitoring the programme. Monitoring activities track the implementation of the programme, and provide answers to such questions as whether it is reaching enough people, or whether the resources are being used as intended. Monitoring data feeds into the management process and into the overall evaluation of a project.

Monitoring may run on continuously in the background, with little or no input from programme delivery staff or participants – for example, very large education programmes may be monitored almost entirely by automated management information systems that track enrolments, drop-outs and contact hours. Alternatively, monitoring activities might include distributing questionnaires to collect information about the types of people participating in the education initiatives.

The evaluation process, meanwhile, is seeking to assess if, and how, the financial education programme is adding value, whether it is meeting its objectives, and what impact it is having. It is considering whether the financial education is leading to a change that would not have occurred otherwise. Evaluation may occur during the programme, so that the results can be fed into the ongoing development of the programme, or it may occur afterwards, to identify important lessons from a programme that can be applied elsewhere.

The reasons for focusing on evaluation

Policy makers and programme designers need monitoring and evaluation evidence in order to be certain that a programme is effective, to identify areas for improvement and to check that the initiative makes good use of resources. Evaluation evidence is vital to inform future funding decisions and to maximise the benefit of financial education provision, but it can also be used as a way of measuring success to build staff morale, improve recruitment and increase awareness in particular programmes.

Monitoring and evaluation can be seen to have both rapid benefits in the short term and longer term benefits. In the short term, monitoring data can be used to check progress and adjust performance so that targets are met. Feedback from participants and interviews with programme staff can lead to beneficial changes in the programme. Over the longer term, the evaluation process can be used to increase the beneficial impact of the programme.

Similarly, evaluation evidence can have a micro level impact, by guiding the improvement of individual programmes and a macro level impact by encouraging the improvement in the overall quality of financial education. A robust evaluation method will provide results that can be generalised to a wider population, so that the relative merits of different approaches can be assessed and the most successful programmes can be replicated and expanded. Furthermore, when robust evaluation findings are generalised to a wider population it becomes possible to predict the overall impact of a programme on a national scale and set well-defined policy targets. Once a number of programmes have been evaluated policymakers can also start to make decisions about the appropriate mix of programmes, the choice of delivery methods and the extent to which other types of intervention may be necessary in order to have the required impact.

The evaluator's role is crucial in making sure that the evaluation provides maximum benefit. The evaluator is responsible for collecting the necessary data and presenting the findings in a clear, unbiased manner. A good evaluation team will draw together the findings from the monitoring and evaluation processes to help programme designers and policy makers to understand where things are working well, what changes might be needed, and what aspects they may prioritise moving forward.

In the following chapter we describe the approaches that have actually been taken to financial education programme evaluation. We also draw conclusions from the stock take of current practice.

Looking at current practice

In the second half of 2009 the OECD INFE Secretariat surveyed all INFE member countries on behalf of the Subgroup on Programme Evaluation to prepare a stock take of the evaluation of financial education programmes and their results.

The country representatives were asked to complete the questionnaire or forward it to appropriate institutions. The questionnaire requested information about the evaluation of financial education programmes, but excluding those delivered exclusively in schools, or the use of national surveys to explore levels of financial literacy, as these are covered by separate INFE subgroups.

Who responded to the stock take?

In total, we received responses from 46 authorities/organisations, representing 29 of the countries that were surveyed[1]. Of these 46 authorities, ten reported that they had not undertaken any evaluation at the time of the questionnaire, although in some cases, evaluation was planned. Eight authorities told us about the evaluation of projects targeting schools and school children; this information has not been taken into account in the rest of this stock take report, but will feed into the INFE Financial Education in Schools subgroup. In total, we are therefore reporting the responses of 28 authorities in 23 countries.

[1] Note that some authorities provided joint submissions and some submitted information about more than one initiative.

In most cases, the authorities told us about the evaluation of financial education programmes that they were responsible for. The exceptions were Finland, Italy, the Republic of Korea and the USA. In the case of the Republic of Korea, Finland and the USA, the reporting authority discussed a number of financial education initiatives that had been evaluated, whilst COVIP, the Italian Pension Funds Regulator, discussed an initiative that was the responsibility of the Ministry of Labour.

What kinds of financial education initiatives have been evaluated?

We asked authorities to describe the financial education programmes in terms of aims and objectives, size, content and delivery methods. Their responses are summarised below.

Aims and objectives

Table 1 Purpose of financial education programme

Purpose	***Number of authorities***	**Percentage of authorities (n=28)**
Raise awareness	*26*	93
Increase knowledge	*23*	82
Change attitudes and behaviour	*20*	71
Raise confidence	*18*	64
Improve financial inclusion	*15*	54

Count of affirmative responses

In total, 11 authorities indicated that their financial education programmes included all of the above objectives, and a further seven left only one objective unmarked. Conversely, just two authorities said that their programme had just one objective.

Size

Table 2 Number of participants

Number of participants reached (or targeted)	***Number of authorities***	**Percentage of authorities (n=28)**
Less than 100	**1*	4
100 to 250	*0*	0
250 to 1000	*6*	21
Over 1000	*12*	43

*Note that the programme with a target of less than 100 was targeting intermediary organisations rather than individuals

Numbers relate to participants actually reached where data available, and targets for newer schemes

Some programmes did not have either a target or a measure of participation, others did not provide responses to this question

Web based initiatives and/or national financial education strategies typically reached the largest number of participants (e.g. New Zealand (c1.5 million), Canada (c2 million), Estonia (c128,000) and the UK (c7.7 million).

The vast majority of the financial education schemes were voluntary, with participants opting to take part. They were largely funded by the public sector (typically by the authority responsible for their delivery). The only exceptions were some of the schemes discussed by the US Federal Reserve Board, the Austrian Financial Driving Licence and the CentiQ programme in the Netherlands, which were part funded by the private sector plus the Belgium INVESTA academy which was the only initiative discussed that did not have Government funding. The soon-to-be-evaluated *Ayo ke bank* initiative in Indonesia is also 80 per cent funded by the private sector.

Content

Most of the financial education initiatives that had been evaluated sought to cover both general financial issues and more targeted issues such as budgeting. Most also attempted to provide participants with improved skills such as being able to make complaints effectively, enhanced financial management, an ability to manage financial risk better or making long-term financial plans.

The content of the initiatives frequently included targeted information on budgeting, managing credit or debt, and savings. These are all aspects of personal finance that have a behavioural element. Relatively complex products such as mortgages, insurances and pensions were also covered in a number of the initiatives discussed.

Only COVIP, the Italian Pension Funds Regulator discussed an initiative that had a specific focus on a single topic (pensions). Amongst the rest, a small minority had a range of content that was targeted towards a specific concern, such as financial difficulties or overindebtedness.

Delivery

Table 3 Delivery methods of financial education programme

Delivery method	***Number of authorities***	**Percentage of authorities (n=28)**
Paper materials	*27*	96
Seminars and lectures	*23*	82
Websites	*23*	82
Special events	*19*	68
Media campaigns	*18*	64

Count of affirmative responses; note that more than one delivery method is possible for each authority

Many of the financial education initiatives used a combination of seminars, paper materials, websites, media campaigns *and* events to reach their target audiences. Conversely, very few of the initiatives that had been evaluated relied on just one delivery mechanism. The exceptions were:

- the Icelandic Confederation of Labour Financial education initiative was delivered solely through seminars, and
- *Pesti Est- Money Compass* in Hungary was delivered via booklets.

All of the initiatives that did not offer seminars or lectures provided paper materials (leaflets and so on). Some also had dedicated websites and/or media campaigns.

The evaluation process

Authorities described the evaluation of financial education programmes in their country in terms of the purpose of the evaluation, implementation, data collection, and the use of benchmarks or control/comparison groups.

Purpose

Table 4 Purpose of evaluation

Purpose	***Number of authorities***	**Percentage of authorities (n=28)**
Satisfaction of participants	*20*	71
Efficiency of programme	*16*	57
Identify necessary improvements to programme	*12*	43
Policy goals	*7*	25
Reinforce accountability	*6*	21
Assess macro economic impact	*6*	21

Count of affirmative responses; note that more than one purpose is possible for each authority

The most common reason for evaluation was to explore the levels of satisfaction amongst participants, yet fewer than half of the evaluations discussed were apparently designed to provide feedback about necessary programme improvements. This may reflect an expectation that satisfaction was likely to be high and that no improvements would be necessary.

Implementation

Table 5 Evaluation implementation

Implementation	***Number of authorities***	**Percentage of authorities (n=28)**
Monitoring included as part of evaluation	*21*	75
Evaluation designed in advance of the programme	*16*	57
External involvement in design or implementation*	*11*	39
Evaluation started in advance of the programme	*3*	11
Evaluation only at the end of the programme	*5*	18

Count of affirmative responses; note that more than one method of implementation is possible for each authority

*Questionnaire asked whether evaluation was designed by the delivery institution, so this is a count of negative responses

Most of the evaluations took place during the programme, regardless of when they had been designed. Five relied entirely on data collected at the end of the programme. Another five indicated that the evaluation included following-up with participants at the end of the programme, or some time later.

Data collection

Table 6 Data collection methods

Methods	***Number of authorities***	**Percentage of authorities (n=28)**
In person feedback questionnaire	*21*	75
Face to face interviews	*10*	36
Web questionnaire	*9*	32
Postal feedback questionnaire	*8*	29
Test	*8*	29
Phone interviews	*7*	25
Market monitoring	*7*	25
Group Interview	*6*	21
Internal data	*4*	14

Count of affirmative responses; note that more than one method is possible for each authority

The stock take indicates that evaluations that were either designed to measure the satisfaction of the participants or to look at programme improvements typically used feedback questionnaires (and, to a lesser extent, web surveys) – hence these being the most popular data collection method.

The efficiency of programmes was more likely to be assessed through tests designed to identify improvements in levels of knowledge, whilst those evaluations looking at the macro impact of an initiative or ensuring that the programme was meeting policy goal drew on national-level data such as levels of complaints about financial service providers.

Some authorities reported that various methods were used to reach as many people as possible during the evaluation. For example, in South Africa, evaluators telephoned those people who were not available for a face to face interview, and posted feedback questionnaires to those who had not responded at the end of the programme.

The stock take suggests that evaluation data has also been collected over very different time periods. ASIC reported that a range of financial education projects in Australia were evaluated over a period of two months, using a mixture of web questionnaires, telephone interviews and feedback forms which took between 10 and 30 minutes to complete. In contrast, the Federal Ministry of Labour, Social Affairs and Consumer Protection in Austria reported on an ongoing evaluation process that aimed to reach all participants with feedback questionnaires, phone interviews and knowledge tests.

In total, 13 evaluations sought to collect evaluation data from all participants. Some used quantitative surveys, others asked open-ended questions on questionnaires, and still others employed qualitative approaches.

Several of the evaluations asked a small number of subjective questions (between seven and 14) to explore satisfaction and self assessment of the effectiveness of financial education programmes.

It seems that whilst many evaluators relied on feedback forms to evaluate courses or seminars, those looking at educational material appeared more likely to use qualitative techniques (sometimes combined with a short data collecting exercise).

The programmes discussed by the Ministry of Employment and the Economy, Finland were unusual in that they were not designed to capture participant satisfaction or effectiveness at an individual level, and have only been evaluated through market indicators and internal data.

Benchmarks and control groups

Table 7 Use of benchmarks and control groups

	Number of authorities	**Percentage of authorities (n=28)**
Benchmark	*10*	36
Control/comparison group	*6*	21

Count of affirmative responses; note that more than one method is possible for each authority

Ten authorities told us that benchmarks existed; these ranged from the results of national surveys through to the goals set by a particular programme in advance of evaluation, or the outcomes identified during previous waves of monitoring and evaluation of the same initiative.

Comparison groups were even more unusual; just six authorities indicated that the evaluations had any kind of comparison group.

What did the evaluations find?

As mentioned above, the various evaluations served different purposes and it is therefore important to recognise that the findings will reflect the purpose of the evaluation and the method used to collect data. The majority of the reported findings focused on the design of the initiative and ways of improving it, since evaluation was most common in that area.

The results of the various evaluations reported in the stock take typically indicated that the financial education initiatives were deemed to be useful and were appreciated by participants. However, some recognised that the initiative could be fine-tuned to better meet the needs of future participants. For example the evaluation of the Belgium *Prevenir et lutter contre le sudendettement en enterprise* scheme indicated that the programme needed reorganising to reduce the intensity of the information given in one day. The evaluation of the Australian financial education offerings highlighted the need to match the programmes to the needs and levels of knowledge of the participants.

Other evaluations found that initiatives were not reaching sufficient numbers of people in the target audience, that intermediaries needed more encouragement to pass on the relevant materials and that levels of confidence amongst participants declined six months after taking the programme. More positive findings included the increase in awareness of the authority in charge (in this case the National Credit

Regulator in South Africa), improved budgeting behaviour, larger proportions of people saving, and increased knowledge. As a word of caution, however, it should be noted that the extent to which these findings can be shown to be causally related to the financial education initiative varies.

Lessons learned

The authorities responding to the stock take were invited to identify the lessons learned from these evaluations, in terms of undertaking further evaluations in the future. Almost all had found the evaluation process useful, and recommended some good practices for other organisations to follow.

These good practices can be summarised as follows:

- Take **independent advice**, learn from other organisations, and share your own results with them
- Undertake **continuous monitoring** of the programme, and analyse the monitoring data regularly
- Include evaluation from the very beginning, planning it alongside the design of the initiative
- Evaluate **regularly**
- Take care to design the evaluation so that it **provides useful information**, and covers the types of people that the programme is intended to reach

Some practical suggestions to improve evaluation processes included:

- Try to automate as much of the data collection as possible
- Identify a suitable comparison group that can be monitored alongside participants
- Be prepared to adapt the evaluation design if it is not meeting its objectives

Organisations had faced a range of challenges in their own attempts to evaluate. These were primarily related to data collection. In particular, the following were mentioned:

- Identifying potential evaluation participants and drawing an appropriate sample (this was felt to be particularly difficult when the evaluation design included a *before* measure, or a comparison group)
- Incentivising programme participants to take part in evaluation
- Undertaking a follow-up survey to identify longer term effects
- Creating questionnaires that were appropriate and relevant
- Ensuring that the responses given were genuine
- Minimising the burden on evaluation participants

- Measuring behaviour change

Other challenges revolved around the cost of evaluation – both in terms of time and money - and ensuring that the evaluation budget was spent to best effect.

Conclusions

This stock take indicates that it is not uncommon for financial education initiatives to be evaluated. There are many authorities undertaking monitoring and evaluation on a regular basis. However, it is striking that each authority has made its own decision about how to evaluate, which methods to use, who to ask for guidance, and what aspects of the education process to focus on. It is clear that a large amount of time and money could be saved if countries could rely on a common framework that allowed them flexibility to evaluate individual schemes using well tested methods. This is a key objective of the OECD INFE Expert Subgroup on programme evaluation.

Many of the challenges that authorities have faced when implementing an evaluation could be reduced or even eliminated by learning from others and sharing common approaches. This current document, along with the two guides developed by the Programme Evaluation Subgroup and the academic review of financial education evaluations are all important steps towards this goal. [2] Furthermore, the development of an internationally comparable survey of financial literacy by the OECD INFE Expert Subgroup and network on the measurement of financial literacy will provide evaluators with a set of questions that can be used to assess levels of financial literacy and look for changes over time.[3] The programme evaluation subgroup may want to go further to provide guidance on practical issues that often occur during evaluations. These may include ensuring the continued participation of learners in the evaluation process, perhaps through the use of incentive mechanisms, or identifying data sources that do not require participants to respond to questionnaires, such as bank statements or credit records.

There is already implicit agreement that certain approaches to evaluation are appropriate for particular delivery methods. For example, qualitative evaluation approaches appear to be the preferred method for evaluations of written materials (whether printed, or online). If there is already a common standard for such things, this could easily be made explicit, so that new evaluations follow the same standard. In this way, authorities can easily look to see what the standard is for programmes that are similar to their own. Of course, this does not mean that programme evaluation should necessarily converge to the simplest design, but by creating consensus it will be possible to compare across evaluations more readily, and improve the overall standard of evaluations through collaborative efforts.

---

2 Guide to Evaluating Financial Education Programs [INFE(2009)15/REV1]; Detailed Guide to Evaluating Financial Education Programs [INFE(2009)15/ADD1/REV1]; A Framework for the Evaluation of Financial Education Programs [INFE(2009)5/REV2].

3 Financial Literacy Measurement Questions and Socio-demographics [INFE(2010)3/REV1].

Chapter 2

A FRAMEWORK FOR EVALUATING FINANCIAL EDUCATION PROGRAMMES

by
Annamaria Lusardi*

It is still relatively unusual for financial education programmes to be evaluated. In this chapter, the main benefits of programme evaluation are discussed, along with the challenges faced by evaluation designers and the resulting limitations of existing evaluations. A five tier evaluation framework is assessed as a potential solution to improve evaluation design whilst still allowing flexibility.

---

* Denit Trust Distinguished Scholar in Economics and Accountancy, George Washington University, USA. This chapter was originally published in OECD (2011), "Improving Financial Education Efficiency: OECD-Bank of Italy Symposium on Financial Literacy". The OECD is grateful to the authors for their valuable contributions. The views expressed here are the sole responsibility of the authors and do not necessarily reflect those of the OECD Committee on Financial Markets and the Insurance and Private Pensions Committee, the OECD Secretariat, the member or non-member countries.

1. The Issue: The necessity of evaluating financial education programmes

Over the past thirty years, individuals have had to become increasingly responsible for their own financial well-being. The shift from defined benefit (DB) to defined contribution (DC) plans in many countries has meant that workers today have to decide both how much they need to save for retirement and how to allocate their pension wealth. Furthermore, financial instruments have become increasingly complex and individuals are presented with new and ever-more-sophisticated financial products. Access to credit is easier than ever before and opportunities to borrow are plentiful. Are individuals well equipped to make financial decisions? Unfortunately, many studies have documented that the majority of individuals lack the knowledge of basic financial concepts, such as interest compounding, inflation, and risk diversification; concepts that form the basis of financial decision-making. Moreover and most importantly, lack of knowledge has been found to be associated with lack of retirement planning, lower wealth accumulation, problems with debt, and poor investment choices (see Lusardi (2008, 2009) for a discussion of these issues).

Perhaps as evidence that financial illiteracy is considered a severe impediment to saving, governments, employers, and not-for-profit organisations have promoted financial education programmes. The effects of financial education programmes have not yet been precisely assessed. Several programmes provide some evidence of a general positive effect of financial education on behaviour, but the impact of specific programmes and teaching methods is still unclear. The question "what works best?" has not been clearly answered. The evaluation of these programmes is critically important.

Programme evaluation is crucial for three reasons: (1) to assess the magnitude of a programme's impacts on participants and the community/population as a whole; (2) to verify how resources and funds are spent; and (3) to ultimately improve the effectiveness of a programme. National governments and private financial education providers need to allocate resources efficiently: resources should go to the programmes that are most effective. They also need to fund programmes adequately and make sure that resources are allocated to the designated objectives. Finally, they need to find ways to improve upon existing programmes. In principle, evaluation is crucial for any type of financial education initiative, irrespective of the size of the programme. Without an evaluation, no programme can claim success. Consequently, proper evaluation should be one of the requirements for obtaining funding for both initial and repeated financial education initiatives.

Currently, not every programme performs an evaluation of its impacts, and when an evaluation is performed, different methodologies are often used. Thus, it is very hard to make consistent comparisons across programmes. Moreover, not all evaluations follow rigorous evaluation methods that allow investigators to pin down the effect due to a programme alone, rather than to other confounding factors. In order to identify the most effective ways to improve financial education, it is important to establish a rigorous evaluation methodology that can be applied to all programmes. As explained by Lyons et al. (2006): "The challenge is to create a tool that is flexible enough to meet the needs of a wide variety of individual programmes, yet standardised enough so that it can be used to make comparisons across programmes." The evaluation should be conducted following scientific conventions in order for results not to be dismissed or undervalued. Moreover, if all evaluations follow similar measurement methods, policy makers will be able to compare the results and gain insights on what is most effective. This evaluation system would give policy makers the opportunity to identify the best methods to make financial education successful and to effectively tailor programmes to specific audiences.

2. Existing Literature: Some difficulties in evaluating financial education programmes

So far, there are no commonly accepted definitions of financial education. The OECD definition of financial education provides a useful framework to refer to:

> Financial education is the process by which financial consumers/investors improve their understanding of financial products and concepts and, through information, instruction and/or objective advice, develop the skills and confidence to become more aware of financial risks and opportunities, to make informed choices, to know where to go for help, and to take other effective actions to improve their financial well-being. [4]

Evaluating financial education has proven to offer many challenges. Several scholars have examined the evaluation of financial education programmes. Reviewed below are several studies in the field, including Fox and Bartholomae (2008), Fox et al. (2005), Lusardi (2004, 2008), Lusardi and Mitchell (2007), Lyons et al. (2006), Hogarth (2006), O'Connell (2009), Crossan (2009), Cakebread (2009), and Collins and O'Rourke (2009). From these studies, the main challenges and limitations of the current state of evaluation studies are identified as follows:

- **Lack of a theoretical framework**. Some programmes target changes in behaviour that are not consistent with what prescriptive models of behaviour would predict. For example, a programme might simply target increases in savings. However, if the target population is young adults facing an upward sloping age-earning profile, economists would argue that they should be borrowing and not saving. Similarly, increasing saving may not be an optimal strategy if participants carry debt. In other words, the outcome of a programme needs to be in line with what is best for participants rather than ad-hoc indicators of behaviour.

- **Potential biases**. There are several potential biases in the evaluation process. The first is self-selection: programme participants are often not chosen randomly. Participants may choose to attend a programme because they are interested in improving their economic situation. In choosing to take advantage of a programme, they may be demonstrating motivation that other individuals do not share. Thus, the effects on this group may overestimate the effects the programme would have on other randomly selected groups of the population. Similarly, some groups may be the target of financial education programmes because of their behaviours, i.e., financial education could be *remedial* and offered to those who save the least or face financial problems. This leads to an underestimate of the effect of the education programme on random groups of the population. Another potential bias is attrition bias: participants may drop out during the programme or may not answer follow-up surveys, causing not only a loss of data but also loss of the representativeness of the sample. Low response rates to surveys conducted months after the programme is also a problem.

---

4 See OECD (2005).

- **Measurement issues**. It is possible that financial education programmes may simply improve how participants report their assets and debt rather than have an effect on saving and debt behaviour. Moreover, because of data confidentiality or lack of access to administrative records, assessment is often based on surveys distributed to participants before and after financial education programmes. However, self-reported information may not always reflect actual behaviour or may measure behaviour with a lot of error.

- **Difficulty proving causation.** Because it is difficult to control for the many factors that affect behaviour, it is difficult to prove that programmes cause changes in behaviour. To do so, it is important to have a rich set of data. Due to the presence of many other factors that are difficult to control for in empirical works, many studies prove correlation rather than causation.

- **Lack of a comparison group**. Most studies do not include a comparison group; as a result, it can be hard to prove that the measured improvements are due to participation in the programme. Comparing programme participants to similar individuals who did not attend the programme provides strong evidence of the changes induced by the initiative. The presence of a comparison group therefore helps isolate the specific effects of the programme. An experimental or quasi-experimental design provides more reliable data than a "descriptive" experiment, in which the changes in participants are evaluated only through a pre-test and a post-test.

- **Lack of comparison of outcome with size of intervention**. Programmes rarely compare outcome with the size of the intervention. For example, some programmes assess the effect of one retirement seminar or the effect of sending participants to a benefits fair. Some of these interventions may simply be too small to generate any effects. If participants have very low financial literacy or face very large search and information costs, one retirement seminar may simply be insufficient to generate any effects. This is not because a financial education programme is ineffective, per se, but because the size and intensity of the programme is insufficient to generate a change in behaviour.

- **Difficulty assessing efficacy of different delivery methods** (e.g., lectures, brochures, and videos). Most financial education programmes combine two or more delivery methods, and few programmes are able to control for individual effects of the different methods. Therefore, it is very hard to determine outcomes of each method and to compare them with similar programmes that use a different number of delivery methods.

- **Practical hurdles**. Most prominent are the high cost of a thorough evaluation, the limited funding for evaluation programmes, and the lack of technical expertise.

- **Publication bias**. Evaluators and financial education advocates may not be willing to publish their studies if the results are unsuccessful. Certain non-experimental designs can allow evaluators to show better programme outcomes than do experimental techniques, consequently biasing evaluation toward that which appears most favourable.

- **Data are often not comparable**. Different studies use different methods, measures, indicators, and parameters, even if they ultimately evaluate the same thing. There is currently no agreement **on the most appropriate indicators,** outcomes, and measurable changes to use in evaluating financial literacy programmes.

Some initial Suggestions on how to improve the evaluation of financial education programmes

The difficulties mentioned above are serious and there has not been much agreement on how to address them. Some of these difficulties can be mitigated via experimental design, standardisation, as is explained below.

The importance of an experimental design

Evaluations following an experimental or quasi-experimental design, rather than merely being a "descriptive study,"[5] offer more reliable proofs of programme effects. An experimental design would use a fully randomised "treatment group" and a "control group." A quasi-experimental design does not randomly select the participants of a programme, but uses a pre-selected group and tries to construct a "counterfactual group" that shares the features of the individuals attending the programme. The counterfactual group, which will not attend the programme, will undergo the same assessment processes and will resemble the treatment group as much as possible in characteristics such as age, sex, social class, and ethnicity. Assessment of the control group and counterfactual group help isolate the effects of the programme from other potential factors influencing the outcome(s), informing the evaluator as to whether the changes in the participants' behaviour were *caused* by the programme or not. As Collins and O'Rourke (2009) remark, a strictly experimental design should be the "gold standard" for an evaluation study because it avoids self-selection bias. However, for practical reasons or budget constraints, it is often not feasible to follow this route. Quasi-experimental studies are easier to conduct and still provide reliable results. Financial education providers who cannot afford expensive experimental evaluations may consider turning to a quasi-experimental design.

The experimental or quasi-experimental design can address the problem of self-selection and the confounding effects resulting from external factors and better assess causality. However, it is important to note that an experimental design alone does not offer solutions for all of the difficulties mentioned above. Although experimental studies are able to control for differences between treatment and control groups, the outcomes of a financial education programme may still be hard to detect if the size of the intervention is not significant enough. For example, a study by Duflo and Saez (2003) undertook an ingenious randomised trial to assess the impact of attending a benefits fair on retirement savings for employees of a major U.S. university. They show that attendance of the fair had a rather modest effect on retirement savings. The study is a good example of a rigorously designed evaluation whose conclusions are hard to interpret: it is not clear whether the small impact on savings is due to the ineffectiveness of providing information and education to employees or to the small size of the intervention, i.e., a single benefits fairs may have minimal impact on behaviour in the face of widespread financial illiteracy and general lack of financial information.

[5] I borrowed this definition from Collins and O'Rourke (2009), who used it to refer to those experiments that simply administer a pre-post survey to assess the outcomes of the program.

Controlling for spillovers

An important issue when conducting randomised evaluations of large-scale financial education initiatives is the spillover of programme impacts from recipients of the programme to non-recipients, a point which has been highlighted in programme evaluation in the field of development economics. Duflo and Saez (2003) show that spillovers happen due to sharing of information and influence people's motivation to increase their financial well-being. In their study, they observe a spillover effect in the form of information transfers between colleagues in the workplace. Spillovers may also influence the measurement of impacts in a larger-scale experiment, such as a financial education project in several communities located in the same geographical region. If the evaluation was conducted as a randomised field experiment in which a specific programme was implemented in some towns and not in others, it is very likely that the programme will find some effects (i.e., the effect will spill over) for the communities that were not "treated." Positive spillover effects are beneficial overall as they help financial education providers indirectly reach a larger number of people. Spillover effects are a potential problem for evaluators, however, because they risk diminishment of the measured impacts of a programme: If the control group of a randomised-trial evaluation benefits from a programme attended by the treatment group, a comparison of the two groups will be less likely to show any impact. The evaluation conducted by Duflo and Saez (2003) was able to measure the degree of information transfer between employees of a U.S. university.

For programmes implemented on a larger scale (e.g., many schools or communities in an entire region), the field of development economics offers insights on how to get around these problems or to measure the spillover itself. Miguel and Kremer (2004) measure the spillover effect of deworming treatments in primary schools in Kenya by randomly selecting a third of the schools in a region. By comparing attendance rates in primary schools, they found that non-treated schools located closer to treated schools also had increased attendance rates after the administration of the treatment. Furthermore, by treating only a portion of the student body in a school, they showed that non-treated pupils in treated schools also became healthier, simply by interacting with treated children.

Programme design and evaluation design

The planning of a programme evaluation should begin with the planning of the financial education programme, should receive a budget consistent with its objectives, and should be tailored to the characteristics and the specific objectives of the initiative. Programme implementation and evaluation go hand in hand; the evaluation should be part of the programme itself and not done as an afterthought. Only in this way can pre-implementation, baseline data be collected.

Standardisation and flexibility

Financial education programmes are very diverse. They embrace a wide range of topics, objectives, audiences, and pedagogic methods. For this reason, evaluation studies should be flexible enough to be tailored to the different programmes while maintaining a determinate standard. Such standardised evaluations would allow implementers and policy makers to compare programmes and answer important policy questions:

- What and of what magnitude are the impacts of financial education programmes on the financial literacy level of the participants?

- What types of programmes are most effective? (school-based programmes, after-school programmes, workplace programmes, etc.).

- What delivery methods are most effective? (counselling, workshops, lectures, interactive exercises, etc.).

- Is financial education the most effective way to improve financial literacy? Are there other initiatives that can achieve better results?

- Which programmes should the government implement and which should it discourage?

- Given the specificity of each programme and the necessity for comparison of results, there is often a need for a compromise between flexibility and standardisation, perfect tailoring and homogeneity.

Privacy Issues and respect of confidentiality

One important concern shared by many programmes is the treatment of confidential data and respect for privacy. This is important to improve response rate, moderate attrition biases, and address some of the measurement problems discussed above. Evaluation questions may ask participants to report their annual incomes, divulge their income sources, or provide other information regarding their financial situation (e.g., whether they have a savings account and how much of their income they allocate to a pension fund). Participants may feel uncomfortable sharing this information with evaluators, particularly if they do not know who will be able to access the information they provide. Fear of sharing confidential data may significantly decrease the response rate. For example, participants who are not citizens may be afraid to provide information regarding their sources of income, as they worry they could be reported to governmental authorities for irregular labour activities. At small retirement seminars at a single firm, employees may fear possible repercussions of sharing private financial information with their employers.

Evaluators should always respect confidentiality when conducting tests, focus groups, and follow-up surveys. Moreover, they should explain privacy policies to participants so those participants can complete the evaluation more confidently. There are several best practices to assure the respect of privacy of financial information. A primary confidentiality measure is to destroy all data, including contact information, in a timely and secure manner after the conclusion of the evaluation. In this way, participants can be sure that once they have filled out all necessary evaluation forms, including a possible follow-up survey via mail, their contact information will not be passed to anyone else. Another important practice is creation of a system of ID numbers connected to participants. The participants use ID numbers instead of names on every evaluation form they fill out, and the evaluator is not able to track a form to a specific person (see NEFE online evaluation toolkit by Lyons, Jayaratne, and Palmer for more information on the ID number practice).

Some ways to facilitate the evaluation of financial education programmes

A national Benchmark and a national strategy for financial education

Establishing a benchmark for the financial literacy level of a country can be crucial to facilitating financial education programmes and their evaluation; a national survey measuring financial literacy and financial behaviour is useful for the design and evaluation of both large-scale and small-scale financial education initiatives. A national survey can provide key insights into the state of financial knowledge and the demographic groups that are most lacking in that knowledge. Financial educations providers can then use this information to tailor programmes to the needs and characteristics of the targeted population. For evaluation purposes, a national survey can establish a baseline financial literacy level that can be used as a yardstick in assessing the effectiveness of a financial education programme. The design and evaluation of financial education programmes of any size, from local seminars to large public policies, can greatly benefit from the existence of a national benchmark.

A national strategy for financial literacy could also provide critical help for evaluation as follows:

- identify the main areas where intervention is needed (e.g., access to credit, default in loan payments, lack of information);
- determine the at-risk populations (e.g., young people, single parents, senior citizens);
- give general guidelines for smaller financial education providers; and
- identify the organisation that can assume a leading role in the financial education effort in the country, coordinating the work of financial education providers, and minimising the overlaps.

Agreement on Indicators and establishment of an international benchmark

As Crossan (2009) points out, there currently are "no proven or agreed indicators or measures for financial literacy." Such an observation calls for the need to create a standardised, possibly international, indicator of financial literacy. This would have several advantages:

- It would allow the impacts of one programme on financial literacy to be compared to those of another programme nationally and/or internationally.
- It would allow for the establishment of a national benchmark, an average financial literacy score for each country, utilising the same measures that evaluate financial education programmes, so that:

  a) nations could compare their financial literacy levels with those of other countries;

  b) each programme could compare its outcomes with the national benchmark score; and

  c) national programmes could track their impacts over the years.

The establishment of a common measure of financial literacy and a national benchmark facilitate not only the evaluation of financial education programmes but also the comparison of the effects of these programmes across countries.

3. A five-tier evaluation framework

To address the limitations of current evaluation studies, several scholars have proposed the creation of an evaluation framework that could serve as a guideline for all programme implementers and evaluators. Such a framework could help the implementers design the evaluation in a standardised way while tailoring it to their specific programme. In *Evaluating the effectiveness of financial education programmes* (2009), O'Connell proposed a new version of a five-tier framework that was first introduced by Jacobs (1988) and later elaborated upon by Fox and Bartholomae (2008).[6]

O'Connell's five-tier approach has many advantages and constitutes an important step toward a standardised but flexible scientific evaluation. In fact, it allows researchers and evaluators to follow an experimental design if desired, to tailor the evaluation to specific audiences and objectives, and to maintain a large degree of similarity with the evaluations of related programmes. The framework is rather broad and provides overall a general direction toward which evaluators can work. More discussion about this approach is reported below.

O'Connell's denomination of the five tiers is different than proposed by others; Le Brun (2009) proposes a more conventional denomination used by scholars in the field of evaluations: the so-called Traditional Model Approach. Despite the new denomination, the content and function of the tiers remain the same. Table 1 compares the two denominations.

Table 2.1: The two denominations of the five-tiered evaluation framework

O'Connell denomination		Traditional Model
Needs		Programme Objectives
Accountability		Programme Inputs
Fine-Tuning		Programme Delivery
Micro Impacts		Programme Outcomes
Macro Impact		Programme Impact

As previously mentioned, programme evaluation should go hand in hand with programme design and implementation. As Fox and Bartholomae (2008) point out, the five-tiered approach "encourages evaluation to occur in each stage of programming."

There are three overarching themes of the evaluation that cut across the tiers. As also discussed by Hogarth (2006), these themes are: 1) objectives, 2) audience, and 3) available resources. They are crucial not only for the design of the evaluation, but also for the planning of the programme itself.

6 See O'Connell (2009) for detailed explanation of this framework.

Objectives: The questions to answer are: What are the objectives of the programme? What aspects of financial literacy is the programme trying to improve? This theme is not only part of the Programme Objectives tier—where the objectives are defined—but affects all the other tiers as well. In fact, the implementer should always consider the programme objectives when identifying the input (what type of class, how many hours, what budget, how many teachers, etc.), the delivery system (what type of teaching system best suits the objectives: lecturing, interactive lecture, activities, workshops, online course, personal counselling, etc.), and the outcomes (based on the objectives that were identified, which outcomes should be expected: e.g., better control of financial transactions and bookkeeping if the objective of the programme is increasing responsibility for and care of household finances).

Audience: The questions to answer are: What type of audience is the programme targeting and who are the programme participants? These issues are first addressed in the Programme Objectives tier, as the objectives may be strictly connected to a specific group of people, but they should then be considered in all the other tiers. When designing the programme, the implementers should ask what types of input and delivery systems best suit the audience (e.g., a primary school programme would be more likely to reach a large number of students, with several hours of classes distributed throughout the year, and use interactive exercises and activities to help children learn; on the other hand, a retirement workshop for a firm's employees might aim at two or three meetings in the same week, with a restricted number of participants, and have a more straightforward, lecture-type delivery method). The outcomes and impacts will also be measured in different ways depending on the audience. For example, high school students might receive surveys that assess critical thinking and direct application of the knowledge acquired, while adult employees might receive a straightforward survey assessing knowledge, attitudes, and behaviour.

Available resources: Evaluations are costly. Some methods are more expensive than others; given the budget constraints of most programmes, not all methods will be affordable. Evaluators should carefully consider existing resources and budget constraints. Thus, objectives, inputs, and delivery methods, for example, have to be chosen in keeping with the resources that are available for the programme and the evaluation.

4. Implementation of the five-tier framework

O'Connell's five-tier approach is a sound basis on which to create a more detailed framework for the evaluation of financial education programmes. The tiers give good direction for evaluators, but more specific guidelines and information on tools, measurement methods, and indicators are needed. This chapter provides further implementation of the five-tier framework to establish a more detailed guideline for future financial education evaluations. Each tier is analysed, giving a quick definition of its objectives, and possible options are provided of what could be measured in each section and how it could be measured. For the last two tiers, where evaluation could take very different paths depending on the type of programme, some of the limitations of the potential methods are analysed. Le Brun's (2009) denomination is used but also O'Connell's (2009) tiers is included in Table 2.

Table 2.2: The implementation of the five-tiered framework

O'Connell Denomination	Traditional Approach Model	What to Measure	How to Measure It
Need	Programme Objectives	Financial literacy scores Levels of bankruptcy High debt Low savings	Financial literacy survey to target group Existing financial literacy results Publicly published data on specific issues
Accountability	Programme Inputs	Costs of the programme Duration of the programme Size and characteristics of the target group Teaching methods/delivery	Description of teaching methods, etc. Survey at the start or end of programme Data on programme expenses
Fine-Tuning	Programme Delivery	What in the programme was effective What was not effective What can be improved What should be changed or completely eliminated	Post-survey to participants and teachers Focus groups One-on-one interview
Micro-Impact	Programme Outcomes	Knowledge Outcome	Pre- and post-test Alternative assessment Criterion reference group test Follow-up survey
		Attitude Outcome	Pre- and post-test Alternative assessment Follow-up survey Focus groups – One-on-one interview
		Behaviour Outcome	Follow-up survey Alternative assessment Focus Groups – one-on-one interview Administrative records
		Other Outcome	Follow-up survey
Macro-Impact	Programme Impact	Financial literacy scores	National survey conducted regularly
		Other indicators of financial well-being: i.e., income level, savings level, indebtedness.	National survey conducted regularly

5. Analysis of the five tiers

Programme objectives

Programme Objectives is the tier that deals with the purpose of the programme and identifies its aims. The objectives are normally addressing one or more problems related to financial literacy, which could range from general lack of financial literacy to a more specific issue such as high number of defaults on mortgage payments. The evaluator should verify the existence and relevance of such issues before the programme is implemented. O'Connell called this tier "Needs" because it identifies what should be improved in the community. In order to identify needs and objectives, the evaluator should also identify the target audience (e.g., high school students, retiring workers, low-income families).

Below are some of the possible measures to verify needs and set objectives that Fox and Bartholomae (2008) provide:

- Testing financial literacy levels in the target population in the case of a general financial education programme. The implementer can take advantage of pre-published data if they exist; otherwise financial literacy tests can be administered among a random sample of the target group.

- High rates of small-business bankruptcy, high levels of consumer debt, and low savings can be used as indicators of poor financial literacy and of poor financial management. To evaluate needs using these indicators, the implementer should find statistics and data on the target community.

- More specific data can be used for particular programmes, such as high levels of default in the repayment of mortgages for a programme on financial management for families that are planning to buy a house. Publicly available data or new survey data could be used for such issues.

Programme inputs

This tier analyses the "inputs" of the programme. It deals with the collection of information regarding the education programme itself and the service provided: its costs, its length, what it teaches, and who participates (number of participants and particular characteristics such as ethnicity, education level, and income level). As Fox and Bartholomae (2008) discuss, the goal of this tier is to assess whether the target group is affected by the programme and in what way. Precise information about the target group is also important to create a well-tailored control group.

Information on inputs can be gathered with a survey given during programme registration, at the end of the programme, or at another appropriate time.

Any programme should pay attention to the inputs both for cost considerations and for accountability.

Programme delivery

The programme delivery tier assesses the way the programme was implemented, whether and what was effective, and what should be changed to improve it. One effective way to receive feedback on programme delivery is by directly asking both participants and instructors. This can be done through specific surveys at the end of the programme. Students and instructors should be provided with different surveys, as their observations on the programme come from different perspectives. The survey could ask participants and instructors to rate the effectiveness of specific teaching methods or course material as well as ask open-ended questions (e.g., what was the most/least effective part of the course?). While participants could be asked how the programme affected them, the instructors could be asked what improvements they saw in participants. Other possible evaluation tools in this tier are focus groups and in-depth interviews with participants or instructors. These two methods can provide valuable qualitative data to better put into context the results of the quantitative data collected through the survey. They might substitute for open-ended written questions and allow for additional comments.

Programme outcomes

All direct effects of a programme on participants are considered programme outcomes. Outcomes can generally be divided into subsets: knowledge outcomes, attitude outcomes, behaviour outcomes, and practical outcomes. Each subset assesses different aspects of the overall effect of the programme and is measured with different, although often overlapping, methods.

Knowledge outcomes

Knowledge outcomes measure changes in programme participants' knowledge of specific financial concepts (e.g., how compound interest works) or programme-related concepts (e.g., how a firm's retirement plan works).

There are several methods through which to assess knowledge outcomes:

- **Pre- and post-test**. This is a test or survey administered to participants before and after a programme to measure change in knowledge of general financial literacy concepts and/or concepts specific to a programme. The pre- and post-test should be well tailored to the programme audience and objectives. The design of the pre- and post-test must account for age and education level of participants. For example, school children who participate in a course on financial responsibility could receive a test with very simple language, with the questions perhaps supported by images. Low-education employees could be given a test with language and concepts that are simpler than those directed to employees who are college graduates. The results of a post-programme test can be compared both with pre-programme and control group results.

- **Criterion Reference Group Test** (CRT). This test can be used for specific groups of individuals who may have problems with reading and writing (for example, very young or very old participants, or participants with low education). The CRT is similar to a group pre- and post-test, but is conducted orally. The instructor asks the questions out loud to the group; individuals may answer and other participants can supplement or correct those answers. Based on correctness and thoroughness, scores are assigned to the answers. This type of evaluation is also applicable to groups who are unlikely to be familiar with or who are uncomfortable with written tests.

- **Alternative assessment**. The alternative assessment presents participants with hypothetical scenarios, vignettes, and exercises to which they have to apply the concepts they have learned. The alternative assessment is mainly used to assess attitudes and expected behavioural changes rather than objective knowledge or actual behavioural changes. The alternative assessment has both advantages and limitations.

Follow-up survey. A survey conducted some time (months or even years) after the programme can measure knowledge retained by individuals long after their participation. The evaluator should be aware, however, of the many external factors that might contribute to and influence the knowledge of programme participants in the intervening time period.

Attitude outcomes

Attitude outcomes measure the effects of a programme on participants' attitudes toward financial literacy and financial responsibilities (e.g., the participant plans to keep closer track of his/her expenses in the future, says he/she will seek more information before purchasing a new financial product) or toward the topic specifically dealt with by a programme (e.g., the participant now believes it is very important to plan ahead for retirement). Attitude outcomes also assess the participant's level of self-confidence regarding general or specific issues (e.g., the participant feels more/less confident dealing with loans and mortgages) following the programme.

There are several methods through which to assess attitude outcomes:

- **Pre- and post-survey**. The same pre- and post-test used for knowledge outcomes can include a survey to assess attitudes and self-confidence. The attitude results before and after the survey can be compared; participant responses could also be compared with those of a control group.

- **Alternative assessment.** The alternative assessment exercises used to evaluate knowledge can also be used to assess the self-confidence of participants in dealing with particular issues or handling specific concepts.

- **Follow-up survey**. Attitudes can be measured months or years after the programme via a follow-up survey.

- **Focus groups**. These groups are selected to discuss and share their experiences, attitude changes, and actions they are planning to take in the future. Focus groups should normally be held immediately after the conclusion of a programme. Focus groups scheduled for weeks or months after an initiative are less likely to be informative, as the direct effects of a programme may be disturbed by other external factors.

Behaviour outcomes

Behaviour outcomes are changes in behaviour normally associated with general economic issues (e.g., the participant has opened a savings account and/or keeps better track of his/her transactions) or specific issues addressed by a programme (e.g., the participant has changed his/her retirement plan or has bought a different financial product). Behaviour changes can be measured months or even years after a programme.

One way to assess the impact of a programme on behaviour is with follow-up surveys conducted a determinate amount of time after a programme. Done via phone or mail, or even face-to-face, such surveys assess how participants' financial behaviour has changed compared to earlier behaviours, which were identified through a pre-programme survey. Focus groups can also be used to gather qualitative data, asking participants to discuss their changes in financial behaviour. Another method is to collect data on behaviour using administrative records, for example bank records, employer records, or government records.

Other outcomes

Other Outcomes refer to specific outcomes, including changes in the financial well-being of participants. A follow-up survey can be used to determine these outcomes. Participants can be asked questions about their level of savings, investment income, retirement plans, satisfaction with their financial situation, or other indicators of financial well-being. The results should then be compared with the data collected via a pre-programme survey and with national data, if they exist. Countries can differ widely on measures of financial well-being and these differences need to be taken into account in performing an international comparison.

Programme impact

This last tier refers to the programme impact at the macro level and is relevant only for programmes that have a large scope (national or at least regional). The issue addressed in this tier is the effect of the programme on society as a whole; its impact relative to other possible initiatives to increase financial literacy and financial well-being in general. The measurements of this tier will be statistics and data at the national level regarding, for example, financial literacy, level of savings, percentage of households with checking/savings accounts, overall self-confidence of the citizens, etc.

This last tier reinforces the importance of establishing a national benchmark and keeping track of changes over time in financial literacy and other indicators. The evaluators in this case need to assess whether programmes (such as mandatory school-based financial education) create improvements over time in the entire community, not only among the people who participated in a financial education initiative.

One way to evaluate macro-effects is by conducting national surveys to analyse trends and changes in financial literacy and in other indicators. In order to do so, a standard survey with a determinate set of questions and a common measurement method should be established and maintained without alteration. In this way, the results of the surveys can be compared over time. These surveys can be conducted by the government or other agencies.

Some pitfalls in the implementation of the five-tier approach

The implementation of the five-tier approach faces many challenges. Some of these problems are again highlighted below and they can instrumental in the success or failure of the implementation of the programme.

Programme objectives

Sometimes the objective of a programme is not properly spelled out or is not well known to the evaluator or the reviewer of the programme. For example, some financial education programmes may be initiated to satisfy specific regulatory restrictions. Employers may offer retirement seminars to comply with laws. Similarly, financial counselling has become mandatory in some bankruptcy procedures. While the improvement of financial literacy may be a declared objective, it may in fact be secondary to the objective of complying with the law. Several studies have been done, for example, to assess the impact of employer-provided financial education programmes in the United States. However, the investigator often had limited or no data on the reason for initiation of the financial education programmes.

The objectives of the programmes also call into play the importance of an independent evaluation. Because the objectives of programmes are often improvements to financial literacy or to financial behaviour, there is a potential bias in reporting only the parts of the programme that work and downplaying or not reporting what does not work. However, the latter can be of great importance, too. Similarly, there is an incentive to choose evaluations that are flexible and tend to favour finding an effect, such as relying on descriptive methods rather than experimental methods.

Programme inputs

Very few studies report the costs of implementing a programme and it is consequently hard to truly assess effectiveness. Moreover, it is hard to make comparisons across programmes. One way to do so would be to rely on indicators such as the return on investment, which are, however, not applicable to every programme.

Programme delivery

Many programmes do not use more than one delivery method and it is therefore hard to disentangle whether (lack of) effectiveness is due to the programme design overall or simply to its delivery. For example, programmes relying on brochures, calculators, and heavy statistics may be unappealing to participants with low financial literacy. The programme may in fact be effective with a different audience but a mismatch between delivery method and recipient characteristics can make it ineffective.

Programme outcomes

As mentioned before, it is difficult to measure the outcome of the programme properly.

Pre-test, post-test, and survey

The main weakness of the pre- and post-test is that some of the indicators of attitude, self-confidence, and behaviour are self-reported. They can be biased indicators. The participant may also be uncomfortable reporting data on savings and debt and/or report them with error.

Follow-up survey

A follow-up survey is a good method through which to assess changes in behaviour, knowledge, attitude, and well-being. However, it faces challenges and limitations. One hurdle is **cost**. This type of survey, conducted via phone, mail, or face-to-face interview several months or years after the programme, is often very expensive.

The second limitation is the so-called attrition bias: as noted by Collins and O'Rourke (2009), many participants will not reply to a mail survey or will not be reachable by phone. This will cause a significant loss of data. There are other biases as well: the individuals who respond to a follow-up survey may be those who are more motivated to improve their financial well-being. Therefore, the results may be biased toward finding a result. Another possible bias is the fact that data on attitude, income, and financial well-being are self-reported and the evaluator often cannot prove whether the interviewed person is answering truthfully. Another limitation is that a follow-up survey cannot in any way control for other variables that might affect the results. For example, participants might have taken more financial education courses in the time period between programme and follow-up survey; they might have received a promotion at their workplace, etc. These events can affect behaviour and are hard to control.

Focus groups

Focus groups are a good method through which to obtain qualitative data on participants' opinions on the effects of a programme. They give the evaluator an opportunity to collect extended data that closed-ended questions and numerical scores of the pre- and post-test are often not able to communicate. With focus groups, specific details of the programme can be discussed in detail, including complaints and suggestions for improvement.

Programme impact

Delays in publishing data

One major limitation of assessing the impact of large programmes is the long time frame necessary to collect, elaborate, and publish data from large surveys. It may take months or sometimes years to collect data at the national level and to elaborate the results from the data. The costs of such surveys are high and often require a significant amount of resources.

Baseline and national surveys

Without a baseline to refer to in order to assess the impact of a programme, every single programme will have to collect data. Because baselines can be useful for a variety of programmes, it may be particularly valuable to centralise the collection of those data.

6. Recommendations

There are several recommendations that emerge from the five-tier approach, particularly for private and not-for-profit institutions. These institutions play a key role in promoting financial literacy, given the importance of a grassroots approach to promote financial literacy and financial education programmes.

The primary recommendation is simply a reinforcement of the importance of evaluation. Without an evaluation, no programme can be considered effective, and this can severely limit not only its adoption by other institutions, but also its funding. Thus, time, effort, and resources need to be allocated not only to programme design but also to programme evaluation. And, as mentioned throughout the report, design and evaluation should go hand in hand, and the evaluation should be part of the programme from the very beginning.

Given the inherent difficulties of measuring the effects and assessing the impacts of a programme, a variety of methods and tools should be employed. Using different methods will allow evaluators to obtain a more complete and multifaceted view of a programme's impacts. For example, both qualitative and quantitative data can be part of an evaluation. They serve different purposes and can provide complementary insights into the impact of and ways to improve upon a programme.

Given the many biases in evaluating the effectiveness of financial education programmes, experimental or quasi-experimental methods should be given priority. It is often hard - if not impossible - to get around self-selection biases and be able to assess the causality between financial education and financial behaviour. Being able to rely on a control and a treatment group both facilitates the assessment of the impact of the programme and gives the programme more credibility.

Clearly, one of the best ways to evaluate an educational initiative is to enlist an independent agency with professional expertise in the field to conduct the evaluation. However, professional agencies are expensive and their costs can be prohibitive for small, private organisations. Public and private higher-education institutions (such as colleges and universities) are perhaps an alternative to professional evaluation agencies. Higher-education institutions are frequently interested in partnerships which give them the opportunity to do research and gather new data. Small organisations could also take advantage of free (online) evaluation toolkits and resources, for example those provided by the OECD, their national governments, educational institutions, or other private organisations. Online resources provide important material that can help prevent common mistakes and propose best practices. Toolkits like the one designed by Lyons, Jayaratne, and Palmer for NEFE are user friendly and flexible, giving evaluators the ability to use established questions on a wide range of topics or create their own questions while maintaining the standardised layout of the evaluation.

Continuing on the previous point, evaluators should rely on external support (and partnerships) to conduct their analysis. Peer-reviewing of the evaluation methodologies before the implementation of the study and prior to the publication of the results is highly recommended.

Resources dedicated to financial education and its evaluation are scarce. Some coordination may be not only helpful but will likely benefit the community at large. For example, a baseline establishing the level of financial literacy and measuring at-risk financial behaviour can be done in a centralised way, without having each institution run a survey. Moreover, there may be a lot of overlap in the type of programmes and the target population. For example, many programmes have set up websites to provide information and help with financial decisions and, in doing so, institutions may end up replicating the efforts of other similar organisations.

There has been little attention to dissemination of results in the discussion of programme evaluation. However, one of the objectives of an evaluation is to prove its significance in order to be compared with other programmes or be adopted at large. Therefore, not only should the results of the evaluation be promptly made available but the evaluator should describe in detail the methodology used for the evaluation. A precise explanation of the methods is crucial to allow for their replication in other studies. If feasible, the evaluator should benchmark the outcomes of the programme with the results of earlier studies on similar initiatives. All results should be reported, not only those that provide evidence of a positive impact of the programme on behaviour. Knowing what is not effective can be as important as knowing what is effective. Results should be made available to any interested parties. For example, an evaluation study could be uploaded to the website of the organisation or agency that conducted the programme. The organisation could also share its report of the initiative with local, national, or international clearinghouses for financial education, such as the IGFE. The evaluation study should reach as many people as possible in order to share with other agencies the findings of the initiatives and allow other financial education providers and evaluators to build on previous knowledge and experiences.

Finally, the evaluation must follow proper guidelines to protect the rights of participants. Data confidentiality, proper disclosure, and securing privacy are not only necessary requirements but are also important requisites to limit attrition biases. Also, participation in the programme has to remain voluntary and evaluators have to think hard about the potential ethical issues of treating groups differently.

7. Conclusions

In this chapter a five-tier framework has been presented and discussed that is directly applicable to different types of financial education programmes. The framework is a simple guideline that financial educators can follow when designing evaluation studies. While leaving significant flexibility to the evaluators, the framework provides a high degree of standardisation, which will allow programmes to be compared both nationally and internationally.

Evaluators have the responsibility to follow the framework and apply it to their programmes in the most effective way. Their diligence in doing so will lead to progressively deeper insights into which methods are the most effective in assessing the impacts of financial education programmes, and perhaps even lead to the creation of novel evaluation methods. Therefore, the five-tier framework is just the starting point, a compass for policy makers and educators in the field.

The debate about financial education programmes and evaluation methods is far from over. The continued discussion and interaction among scholars, evaluators, and policy makers is necessary to enrich and improve upon the existing evaluation studies. Further efforts are needed to apply the five-tier framework to as many programmes as possible in order to assess the effectiveness of different initiatives and collect information on the needs of specific groups of citizens. In this way policy makers and private organisations will be able to tailor financial initiatives to specific needs, using the most effective methods. And the OECD can become an important resource by creating a supporting structure for all policy makers and financial educators.

The interest raised by financial education initiatives in many countries underlines the importance of financial literacy and its link to financial well-being. Financial education is increasingly becoming a priority among policy makers and private institutions in countries around the world. It is therefore important for institutions like the OECD to facilitate discussion and sharing of ideas among its members.

References

Cakebread, C. (2009), “Is Financial Education in Canada Working? Research and recommended best practices for evaluating financial education programs.” Working paper for the Financial Consumer Agency of Canada (FCAC).

Caliber Associates (2005), “JA Worldwide: JA Dollars and $ense pilot program formative evaluation, final report.” For Junior Achievement Worldwide. http://www.ja.org/files/afterschool/JA_Dollars_&_$ense_Final_Formative_Eval_Rpt.pdf

Collins, J. M., O’Rourke, C. (2009), “Still holding out promise: a review of financial literacy education and financial counselling studies” a report for the Annie E. Casey Foundation.

Crossan, D. (2009), “Evaluating financial education programmes” for the Retirement Commission, New Zealand.

Duflo, E. and Kremer, M. (2008), “Use of Randomization in the Evaluation of Development Effectiveness” in Easterly, W. (ed.) *Reinventing Foreign Aid,* Cambridge, MA: The MIT Press, pp. 93–120

Duflo, E., and E. Saez (2003), “The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment.” *Quarterly Journal of Economics*, 118, pp. 815–842.

Fox, J. J., and Bartholomae S. (2008), “Financial education and program evaluation” in Xiao J. J. (ed.) *Handbook of consumer finance research*. New York: Springer, pp. 47–68.

Fox, J. J., Bartholomae S., and Jinkook L. (2005), “Building the case for financial education.” *The Journal of Consumer Affairs*. Vol.39, No.1, 195–214.

Glazerman, S., Levy, D., and Meyers, D., (2002), “Nonexperimental Replications of Social Experiments: A Systematic Review.” *Mathematica Policy Research, Inc*. http://www.mathematica-mpr.com/publications/PDFs/nonexperimentalreps.pdf

Hogarth, J. M., Federal Reserve Board, USA (2006), “Financial education and economic development”. Prepared for *Improving Financial Literacy: International Conference hosted by the Russian G8 Presidency in Cooperation with the OECD*.

Jacobs, F. H. (1988), “The five-tiered approach to evaluation: Context and implementation” in H. B. Weiss & F. H. Jacobs (Eds.), *Evaluating family programs*, New York: Aldine DeGruyter.

Le Brun, M. (2009), “Analysis of O’Connell framework for evaluating financial education programmes.”

Lusardi, A. (2004), “Saving and the Effectiveness of Financial Education,” in O. Mitchell and S. Utkus (eds.), *Pension Design and Structure. New Lessons from Behavioral Finance*, New York: Oxford University press, pp. 157–184.

Lusardi, A. (2008), "Introduction," in A. Lusardi (ed). "*Overcoming the Saving Slump: How to Increase the Effectiveness of Financial Education and Saving Programs,* Chicago: University of Chicago Press, pp. 1–13.

Lusardi, A. (2009), "U.S. Household Savings Behaviour: The Role of Financial Literacy, Information and Financial Education Programs," in C. Foote, L Goette, and S. Meier (eds), "*Policymaking Insights from Behavioral Economics,*" Boston: Federal Reserve Bank of Boston Press.

Lusardi, A., P. Keller, and A. Keller (2008), "New Ways to Make People Save: A Social Marketing Approach," in Annamaria Lusardi (ed.), *Overcoming the Saving Slump: How to Increase the Effectiveness of Financial Education and Saving Programs,* Chicago: University of Chicago Press, pp. 209–236.

Lusardi, A., and O. S. Mitchell (2007), "Financial literacy and retirement preparedness: Evidence and implications for financial education." *Business Economics*, pp. 35–44.

Lyons, A. C., Palmer, L., Jayaratne, K. S. U., and Scherpf, E. (2006), "Are we making the grade? A national overview of financial education and program evaluation." *The Journal of Consumer Affairs*, Vol. 40, No.2, p.208–235.

Lyons A. C., Jayaratne, K. S. U., and Palmer L. "Evaluation Manual." NEFE® Financial Education Online Toolkit. (accessed 21/10/2009) http://www.nefe.org/eval/manual_toc.html.

Miguel, E., Kremer, M. (2004), "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica*, Vol. 72, No.1, pp. 159–217.

O'Connell, A. (2009), "Evaluating the effectiveness of financial education programmes" working paper for the Organisation for Economic Cooperation and Development.

Organisation for Economic Cooperation and Development (2005), "Recommendation on Principles and Good Practices for Financial Education and Awareness." http://www.financial-education.org/dataoecd/7/17/35108560.pdf

Organisation for Economic Cooperation and Development (2005), "Improving Financial Literacy: Analysis of Issues and Policies." http://www.oecd.org/document/2/0,3343,en_2649_15251491_35802524_1_1_1_1,00.html#HTO

Pritchett, L. (2002), "It Pays to Be Ignorant: A Simple Political Economy of Rigorous Evaluation" *Journal of Policy Reform*, vol. 5(4), pp. 251–269.

Retirement Commission, New Zealand. "2009 Financial Knowledge Survey." http://financialliteracy.org.nz/sites/default/files/webfm/Research/Financial_Knowledge_Survey_2009_report.pdf

Chapter 3

INFE HIGH-LEVEL PRINCIPLES FOR THE EVALUATION OF FINANCIAL EDUCATION PROGRAMMES

These High-level Principles have been developed by the OECD International Network on Financial Education (INFE) with financial support from the Russian/World Bank/OECD Trust Fund on Financial Literacy, and approved by the OECD Committee on Financial Markets (CMF) and the OECD Insurance and Private Pensions Committee (IPPC) in 2011.

Introduction

In 2003, the OECD launched an international programme on financial education, under the aegis of the OECD Committee on Financial Markets (CMF) and the OECD Insurance and Private Pensions Committee (IPPC). One of the first milestones of the programme was the adoption of the Recommendation on Principles and Good Practices for Financial Education and Awareness by the OECD Council (OECD, 2005).

The OECD Recommendation on Principles and Good Practices for Financial Education and Awareness emphasises the importance of developing efficient financial education programmes, and states that: 'The development of methodologies to assess existing financial education programmes should be promoted' (OECD, 2005). In addition, the CMF, the IPPC and the OECD International Network on Financial Education (INFE) identified the development of guidelines for programme monitoring and evaluation as a top priority. Accordingly, the INFE convened an expert subgroup on the Evaluation of Financial Education Programmes at a global level to encourage policy makers to recognise the value of evaluation and to improve evaluation practice.

Programme monitoring[7] and evaluation[8] are essential in order to be certain that a programme is effective, to identify areas for improvement and to check that the initiative makes good use of resources. Evaluation evidence can also inform national financial education strategies by identifying the most efficient programmes and influencing future funding decisions. Furthermore, when robust evaluation findings are generalised to a wider population it becomes possible to predict the overall impact of a programme on a much larger scale and set well-defined policy targets.

Evaluation data can be complemented by financial literacy measurement at a national or international level. A national survey of financial literacy can help programme designers to identify those topics that should be covered by financial education programmes and to set appropriate targets when writing their aims and objectives. With repeat surveys, it may also be possible to identify changes in levels of financial literacy over time and attribute such changes to the implementation of large-scale financial education programmes. However, a national survey should not be seen as an alternative to programme evaluation, as it does not allow for detailed analysis of the benefits particular programmes and is not designed with reference to the aims and objectives of a particular programme.

Against this backdrop, and with the guidance of various OECD/INFE outputs[9], this document presents the INFE high-level principles for the evaluation of financial education programmes (herein after referred to as the high-level principles) and is intended to complement the 2005 OECD Recommendation. The high-level principles are designed to further inform policy makers, financial education programme designers and other stakeholders about the importance of evaluation. They relate to the monitoring and evaluation of a wide variety of financial education provision, from individual courses and seminars to large scale programmes and media campaigns. The evaluation of national strategies requires evaluation

7 i.e. regular, routine tracking of inputs and outputs.

8 i.e. assessment of changes that can be attributed to the programme.

9 OECD/INFE(2010a), OECD/INFE(2010b), OECD/INFE(2013a).

of the component parts, and these high-level principles are therefore of relevance in all countries regardless of whether there is a national financial literacy strategy.

The evaluation of school programmes is not within the scope of these high-level principles; the OECD and its INFE have developed specific work to address the delivery and evaluation of financial education in schools[10].

Process

These high-level principles were developed and approved by the OECD INFE in December 2010. They were then transmitted to the OECD Committees in charge of Financial Education [the Committee on Financial Markets (CMF) and the Insurance and Private Pensions Committee (IPPC)] and approved for further public consultation. The document was subsequently made available for public consultation on the OECD website and the OECD International Gateway for Financial Education[11] in August/September 2011. Minor comments were received and reviewed through this process. Following the successful public consultation process, the document has been endorsed by the INFE as INFE high-level principles.

1. Evaluation: an essential element of financial education programmes

New programmes should be evaluated

A monitoring and evaluation strategy should be developed alongside new financial education programmes in order to gain maximum benefit from the evaluation data. Evaluation should be given the same importance as any other aspect of the programme. Dialogue and collaboration between programme designers, facilitators and evaluators should be promoted (to the extent possible) in order to make sure that everyone is in agreement about the aims and objectives and the implementation of the evaluation.

Care should be taken to align expectations among stakeholders and to ensure that the evaluation can provide stakeholders with useful information.

Existing programmes should be evaluated to the extent possible

Whilst programmes should ideally be designed with evaluation in mind, there are ways of evaluating existing initiatives in order to provide evidence of efficiency and inform future policy decisions.

All ongoing monitoring processes and available data sources should be identified and used as necessary when evaluating existing programmes in order to avoid duplication of effort. Where no such resources exist, monitoring and evaluation processes should be implemented at the same time.

As with new programmes, the process of designing an evaluation of existing programmes should be collaborative, in order to make sure that all stakeholders are in agreement about the aims and objectives

10 OECD/INFE (2013b).

11 www.financial-education.org

as well as the implementation of the evaluation. Priority should be given to explaining to stakeholders why the evaluation is taking place, and seeking their full cooperation.

2. Budget for evaluation

A budget should be set for evaluation. Where it is not possible to evaluate all aspects of the programme within the budget, stakeholders should be encouraged to identify which aspects of the programme should be evaluated as a priority taking into account the main objectives of the programme.

The amount of money set aside for evaluation should be proportionate to the overall cost of the programme. Particular attention should be given to identifying pre-existing evaluation resources that can reduce the cost of evaluation such as existing evaluation materials, resources or frameworks, including the OECD INFE guides to evaluation[12]. Existing survey data could also be used to set benchmarks; tests or survey instruments can be modified for use in evaluation; and data such as employers' records of money paid into pension funds can be used to identify changes over time.

Whenever possible, consideration should be given to obtaining an accurate estimate of evaluation costs before finalising the funding for financial education programmes in order to set aside an appropriate amount of money or identify additional funding sources when necessary.

There are some circumstances where the long-term benefit of evaluation evidence can justify high cost evaluation in the initial stages. This is particularly the case with pilot studies designed specifically to assess the potential benefit of a new large-scale financial education programme.

3. External evaluators: adding credibility, skills and independence

The use of external evaluators should be encouraged wherever resources permit. Good evaluators have the necessary skills to ensure a robust evaluation design, the contacts to access information that might otherwise be difficult to obtain and the experience to find solutions to any issues that arise during the evaluation process. Their independence can also encourage participants to respond more openly.

When choosing an external evaluator, consideration should be given to whether they have appropriate levels of knowledge about each element of the programme, including the subject matter, the delivery method and the target group. Similar consideration should be given to their skills and experience, including their ability to report the findings of the evaluation in an accessible way.

The share of the evaluation undertaken by an independent expert should depend on the specific nature of the evaluation. When an evaluation is based almost entirely on monitoring data, for example, the external evaluator should review the data and the subsequent analysis to confirm that there has been unbiased reporting.

Effort should be made to encourage doctoral students and others to work in the area of financial literacy programme evaluation in order to increase the number of external evaluators available.

---

12 OECD/INFE (2010a), OECD/INFE (2010b)

4. Appropriate evaluation design

The design of the evaluation should take into account the objectives of the programme, the size and length of the programme, the target audience and the delivery method. It should also be appropriate to the needs of the stakeholders, including those tasked with applying the findings of the evaluation to future programmes. Priority should be given to evaluation that is designed to attribute observed changes amongst the target audience to their participation in the programme[13].

Aligned with well-defined programme objectives

An evaluation should be designed to assess the extent to which a programme has met specific targets and goals and whether it has completed its overall objectives. This is achieved by aligning the evaluation design with the programme objectives.

Financial awareness and information programmes may incorporate elements of three different types of objective, each of which requires a different evaluation design:

- The first type of programme provides consumers with information, and the objective is often to reach as many people as possible. In such cases monitoring data provides sufficient evidence that the programme has met its objectives.

- Other programmes are designed with the objective of increasing knowledge or skills. In these cases, tests should be distributed before and after the programme to indicate whether learning has occurred. Follow up studies can be used to explore the extent to which the knowledge or skill has been retained.

- The third type of programme is intended to develop new behaviours or change existing behaviours. Evaluators should consider employing long-running monitoring and evaluation processes to assess the extent of behaviour change.

Programmes often address more than one of these objectives, and in such cases the evaluation design will incorporate several methods.

If a financial education programme does not have well-defined objectives, evaluators should identify external benchmarks (such as the performance of similar programmes run elsewhere) against which the programme can be assessed. This is more likely to be necessary when a programme is not designed with evaluation in mind.

Appropriate to the size and length of the programme

Financial education programmes should be evaluated regardless of size, but larger programmes provide more options in terms of methods used and the full range of options should be considered in such cases.

---

13 Additional guidance on choosing appropriate evaluation designs is available from OECD/INFE (2013a), OECD/INFE (2010a), and OECD/INFE (2010b).

The evaluation of smaller scale programmes should be designed with particular regard for the implicit budgetary constraints and the small number of participants from which to draw a sample.

Some financial education programmes take place over several sessions, whilst others are delivered at a single point in time. The amount of input expected of participants should be proportionate to the amount of provision received. Financial education programmes designed to run over several sessions can be designed to incorporate repeated participation and should ideally incorporate evaluations that capture changes over time.

Taking into account the target audience and delivery method

When evaluating a programme that is designed to reach the whole population, priority should be given to using national data such as surveys of financial literacy or aggregate economic indicators. Evaluation designs that automate monitoring and evaluating processes should also be considered as cost effective approaches to data collection in such situations.

The design of evaluations that focus on programmes targeting a subset of the population will depend on the extent to which variations in the economic status, education levels, literacy, numeracy or language of the financial education participants may impact on their ability to input into the evaluation process. The extent to which target audiences are transitory or vulnerable must also be factored into the evaluation design.

The design of the evaluation should be aligned with the delivery method in order to contain costs and provide reliable data. For example, users of web-based learning may be too geographically dispersed to interview in person, but could readily be approached by telephone, email or via an online survey.

Attributing change to the programme

Evaluations should indicate to stakeholders the extent to which programmes have changed levels of knowledge, skills, behaviour or attitudes (where these outcomes are in line with the objectives of the programme). Not all evaluation designs can prove a causal relationship between a financial education programme and observed changes; but ideally, priority should be given to designs that can prove such a relationship. The most rigorous way of inferring a causal relationship is to undertake a parallel study with a group of people who did not participate in the programme (a control group or comparison group, depending on how they are chosen). If the programme participants exhibit changes that have not occurred in the parallel study, then it is reasonable to attribute the changes to the programme.

A well designed qualitative study can give valuable insights into the additional factors that may explain any changes observed and should be considered as an important part of evaluations seeking to understand how change occurred. This could include depth interviews, case studies or focus groups.

In order to attribute levels of knowledge or skill to the programme and identify improvements, the evaluation should incorporate a measure of knowledge or skill before and after participation. Longitudinal studies should be considered with repeated follow-up of programme participants in order to identify changes over time and the extent to which such changes are long-lasting.

5. Reporting

Policy makers should encourage the reporting of evaluation findings to the extent possible. Reports should discuss both the positive and negative findings in order to properly inform future programme design.

Evaluations are only truly valuable if people learn from them; they need to learn what has worked, and what has not worked. When evaluations are reported in an unbiased way they can help both the programme designers and others improve aspects of the financial education on offer and avoid costly mistakes in the future. Sharing the findings also allows others to benchmark their own programmes and set targets.

Stakeholders should identify their reporting needs during the design phase of the evaluation. Depending on the complexity of the evaluation design, many months may elapse between data collection and final reporting. During this time, the benefit of interim reporting should be considered, keeping in mind the possibility that findings will change as more data become available.

Selected References

OECD Recommendations available at www.financial-education.org

OECD (2005a) Increasing Financial Literacy: Analysis of Issues and Policies. Paris: OECD

OECD (2005b) Recommendation on Principles and Good Practices for Financial Education and Awareness

OECD/INFE instruments developed through the Russia/WB/OECD Trust Fund for Financial Literacy and Education

OECD/INFE (2010a) Guide to Evaluating Financial Education Programmes

OECD/INFE (2012), High-level Principles on National Strategies for Financial Education.

OECD/INFE (2010b) Detailed Guide to Evaluating Financial Education Programmes

OECD/INFE (2013a), Evaluating Financial Education Programmes: OECD/INFE Stock Take and Framework

OECD/INFE (2013b), Financial Education for Youth and in Schools: OECD/INFE Policy Guidance, Challenges and Case Studies

ANNEX: GUIDE TO EVALUATING FINANCIAL EDUCATION PROGRAMMES*

Who is this guide for?

This guide is for financial education programme managers, educators and stakeholders. It is designed to help you monitor and evaluate your financial education programmes by choosing the best evaluation methods.

What is an evaluation?

An evaluation is the process of gathering information that will help you determine if a programme is making a difference. The right evaluation method will help you to determine a programme's value, impact and effectiveness and take appropriate action. For the best results in evaluating your financial education programme, you need to plan your evaluation at the same time as developing your programme. Evaluation is systematic, evidenced-based, and measures or makes assessments in helping to make decisions.

Why conduct an evaluation?

Evaluating a financial education programme can help develop new insights into what works, assist in rollout from pilot to full programme or expanding delivery, identify areas for programme improvement and help with your applications for funding in the future. It can also be used to disseminate information amongst financial education specialists and other stakeholders.

What kinds of questions do evaluations answer?

Evaluations help to answer questions about financial education programmes that can guide future developments:

- **Did it achieve what was expected**?
  - If so, how were its goals met?
  - What resources were used?
- **Why did some things work** and other things fail?
  - What were the reasons for the successes and failures?
  - What did you learn from the successes and failures?
  - How realistic were its goals?
  - Were there any wrong assumptions from the start?
- **What difference has the programme made**?
  - How did it affect the behaviour, attitudes, knowledge and skills of the target audience, community, participants, programme providers or others?
  - Who benefited from the programme and how and why?
- **How could the programme run differently**?
  - What changes would you make if you were running the programme again i.e. more useful systems, improved knowledge, experienced people?
  - Were the needs of the programme met?
  - Were organisational systems made more effective during the course of the programme; if so how?
  - How could the evaluation findings be useful to future programmes?
- **How does the programme compare with other programmes?**

* Supported by the Russian/World Bank/OECD Trust Fund, this guide was developed by the OECD International Network on Financial Education (INFE) Expert Subgroup on the Evaluation of Financial Education Programmes using a variety of publicly available documents from various countries. It was approved on 18th October 2010 by INFE expert subgroup members and on the 19th October 2010 by INFE members.

The basis of a good evaluation

Plan your evaluation

- Clearly define the evaluation purpose and evaluation questions; choose methods that are appropriate for the purpose and context of the programme and the target audience. In doing this, find out whether an evaluation model has been successful in the past for financial education programmes of a similar background.
- If possible, include an evaluation in the programme design before implementing the financial education programme. This will focus the programme development, align programme and evaluation objectives and facilitate the collection of initial information for your evaluation, such as baseline (initial) measures of financial literacy.
- Clearly define what you consider to be successful and unsuccessful outcomes for the financial education programme, for example, increased rates of financial literacy at the completion of the programme.
- Have a plan for what you will do with the results of the evaluation.

Consider cultural and socio-economic perspectives

- Identify and respect the dignity, cultural/spiritual values and perspectives of all stakeholders (individuals and groups) e.g. different attitudes towards money, savings, investing. Consider different cultural perspectives throughout all stages of evaluation and how will these may influence the type of evaluation you use, the data collected and the way you report and communicate results.
- Take account of the literacy levels and socio-demographic status of the target population when designing and implementing evaluations.

Think about the cost

- Allocate sufficient funds for the evaluation as well as for delivering the programme. You could set guidelines for how much an evaluation should cost as a percentage of the total programme cost, keeping in mind that this will differ depending on the purpose and scope of the evaluation.

Make sure your evaluation is fair, reflective and, preferably, independent

- When possible, the evaluation should be undertaken by an independent evaluation team with appropriate professional qualifications and experience. An independent evaluation team can bring in a new perspective, add value to the evaluation plan through their expertise, have a capacity building benefit and help to strengthen the objectivity of the findings.
- However, if this is not possible, there are ways for stakeholders to conduct evaluations that ensure the results are robust and credible. An evaluation carried out internally may make it harder to objectively analyse the findings and the problems that arise, but by choosing appropriate and proven methods and building in a 'peer/expert-review' component, internal evaluation can achieve similar results.

Choose the right evaluation methods

- Use proven evaluation tools and approaches where they are available.
- Consider identifying an independent expert to peer-review the evaluation methods chosen before the evaluation starts, and allow time to fine-tune the design following their recommendation.

Put your evaluation into place

- Make every effort to minimise interference with programme implementation.
- Identify research participants, and ensure that their participation does not place an undue burden on them.
- Where possible, participation in any programme evaluation should be voluntary and confidential. This can be achieved by following a detailed plan of how to gain informed consent, how privacy and confidentiality will be maintained, how and where data will be collected and stored during the evaluation and the retention and disposal of data once the evaluation is complete.
- Communicate regularly with participants and stakeholders to inform them of progress.

Communicate your results appropriately

- Communicate findings/results in a range of appropriate formats to meet the needs of different stakeholders, including participants.
- Describe the evaluation methods you have chosen in detail so that the approach could be replicated in the future. This will also enable others to consider the robustness of the evaluation findings.
- State any limitations of the findings and threats to their validity. Every evaluation method has these and they need to be pointed out. Be cautious about making broad generalisations that are not supported by, or go beyond, the evaluation evidence.
- Provide useful recommendations and practical ideas for improving the programme.

The evaluation cycle: three key steps

The evaluation cycle involves planning, implementation and reporting. These three steps will overlap and repeat to some extent, depending on the evaluation method used and the length of the evaluation process.

Planning your evaluation

Purpose and scope of your evaluation

- Ensure that you know the programme objectives clearly so that you know what to evaluate. Develop a logic model that describes the connection between the activities of the financial education programme and the expected outcomes and impacts.
- Identify and regularly communicate with all relevant stakeholders about the evaluation process.
- Research the programme's background and whether an evaluation model has been used successfully in the past for similar financial education programmes.
- Talk to key people to determine the purpose, scope, goals and objectives of the evaluation. Work together to identify the questions that the evaluation should answer.
- Have a plan for what you will do with the results.

What kind of evaluation do you need and which design?

- In selecting a design, you should balance the reasons, the resources available and the intended use of the evaluation findings. Decide what kind of evaluation you will need to get your questions answered.
- Choose the kind of evaluation that best suits your purposes given the timeframe for the implementation of the programme and for the evaluation. Identify indicators of success or failure. Determine what methods you will use to collect data and report your findings.
- Develop your tools and methods for collecting information. These should be based around your evaluation questions. Also, plan for a peer/expert review of your evaluation methods.
- Generate evaluation questions based on what you want to know. Consider collecting data from participants, trainers, and others involved in the programme.
- Identify participants and sources of data or other information, such as tests or survey results.
- Create a system for information collection and processing.

Methods and preparation for data collection

- Determine the specific methods of data collection that you want to use that will support your findings.
- Consider using qualitative, quantitative and administrative data.
- Devise a system for the collection of information which is practical, systematic, ongoing, accurate and ethical. Develop relevant tools and pilot these to ensure they work as you intended.
- Participation in the evaluation should be voluntary.
- Select comparison groups or control groups with care and consider how they will be encouraged to participate in the evaluation research.
- Prepare a plan for protecting privacy and confidentiality.

- Ensure informed consent is obtained from participants.
- Plan and inform participants about the use, retention and disposal of information collected.

Implementing your evaluation

Analysing and interpreting data

- Use early results from the evaluation to recommend initial changes to your programme or evaluation.
- If necessary, refine the programme’s logic model.
- Examine information as you collect it (this can save work at the end).
- Identify sources of bias in the results and seek ways of minimising them.
- Integrate findings from multiple methods and sources.
- Identify key patterns in your results, but be cautious about attributing observed change to your programme. Consider other factors that may have been responsible for the change.
- Do not over-generalise your results.
- Acknowledge findings that did not fit the main pattern or and those that did not meet expectations.
- Share findings regularly among your team, including team members who are responsible for other elements of your evaluation process (this will help give you a fresh perspective on the results).

Reporting and using evaluation findings

Reporting results

- At each stage of the evaluation, report back to key stakeholders on progress and any issues faced.
- Understand how the stakeholders want to use the reported results and determine appropriate ways to disseminate final results to them.
- Develop a detailed report providing a description of the evaluation method so that it can be replicated, and include a statement of limitations.
- Be prepared to explain the rationale for the evaluation and be transparent about the process.
- Prepare a summary and conclusion of your findings that is easy to follow and contains useful examples and strategies for future success. Take care to express results clearly and accurately and in a way that key stakeholders will understand.
- Clearly link the findings, conclusions and recommendations in your report.
- Consider making the results of the evaluation available to participants (where appropriate).

Using results

- Ensure that evaluation results are timely and available when decisions need to be made.
- Identify what aspects of the financial education programme worked well, what did not and the reasons why.
- Apply the results of the evaluation to future programme development where possible (this will depend on the design of the evaluation).

Practical data considerations

The evaluation process will involve the collection of various data that can be used to decide the value of the programme or to recommend further improvements. Such data might include:

- written or spoken thoughts and conversations, photographs or drawings. These are very useful for understanding the experiences of people and exploring questions such as why or how something happened. This is called **qualitative data** and can be beneficial when you want to describe the variety of experiences and understand why something happened or didn't happen as expected.

- Data that provide you with numbers, such as counts of people or money amounts. Such data can be analysed to answer questions such as how many people received the education or how much key outcomes changed by. This is called **quantitative data**. This type of data will not reveal the reasons for achieving a programme's objectives. You will need to use qualitative data to explore potential reasons for the strengths and weaknesses of the programme.

Some suggested methods to carry out an evaluation

Evaluation/ research method	What is involved	Benefits	Limitations
Focus groups	• An interviewer asking questions to a small group of people who hear and react to each other's responses	• This provides insight into people's attitudes and views on a particular subject	• Results may be overly influenced by one or two individuals • Does not provide quantitative data • Sample is not designed to be representative
Key informant interviews	• Interviewing experts on a particular subject • Interviews can be structured or free-flowing	• Provides in-depth, expert information • Often cost-effective • Can help to understand complex issues • Helps you to identify how effective the programme is from the perspective of its participants	• Can be expensive • May be difficult to recruit a representative sample of participants • May be influenced by interviewer bias • Do not provide quantitative data
Assessment information e.g. examination	• Gathering information from participants to measure the learning that occurred	• Provides quantitative information that may enable the effects of a programme to be measured	• Does not provide information about how or why effects have occurred • Results may be influenced by factors unrelated to the programme

			• May be difficult to recruit participants who are willing to be assessed
Observations	• Describing the behaviour and actions of participants through direct observation	• Provides information that is helpful for understanding the behaviour of participants	• May be difficult to establish reliability between observers • Subject to observer bias
Surveys	• Gaining anonymous answers to a questionnaire from participants • Can be a survey that is completed privately (such as a postal or web-based survey), or with a researcher present (such as face-to-face or telephone-based)	• Provides immediate results from a large group • Provides a quick view of the attitudes and views of the participants	• May present barriers to participants with lower levels of literacy • May have low response rates • Respondents may not be representative
Administrative information	• Analysing information that already exists prior to the evaluation, such as a register of attendance, financial information, the type and amount of services provided, and demographic information	• Provides excellent number-based results • Useful for identifying trends and patterns • Can be accessed immediately	• Does not provide participant perspectives • May not accurately reflect the 'lived reality' of programme implementation
Diaries	• Participants recording their experiences and what they have learnt about the programme during its course	• Provides small samples of results based on people's views and perspectives	• Does not provide quantitative data • Can be time consuming and expensive to analyse data • The process of keeping the diary can change people's behaviour • May not have complete set of data as participants may stop writing the diary after a few weeks/sessions

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where governments work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

The OECD member countries are: Australia, Austria, Belgium, Canada, Chile, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The European Union takes part in the work of the OECD.

OECD Publishing disseminates widely the results of the Organisation’s statistics gathering and research on economic, social and environmental issues, as well as the conventions, guidelines and standards agreed by its members.

www.oecd.org/finance