Cognitive Interviewing: Impact on non-formal education variables in the Maltese Labour Force Survey

Charlene Abela, Tania Borg and Joslyn Magro

Abstract

The aim of this paper was to assess and improve the quality of the data collected for the non-formal education (COURATT) variable in the Labour Force Survey [LFS] and to address under-reporting issues. A hybrid cognitive interviewing model was applied to test this variable, where probing techniques were used in conjunction to the think-aloud method. (Ryan, Gannon-Slater, & Culbertson, 2012).

For the purpose of this study, a qualitative analysis of the cognitive interviewing data was carried out to reveal respondents difficulties in answering the question. Overall, findings indicate that respondents experienced difficulties in understanding the questions and in retrieving the necessary information to answer it. Moreover, the changes implemented in the questionnaire due to cognitive testing resulted in a break-in-series for non-formal education estimates and a level shift was also recorded for the LLL indicator.

List of Acronyms

CASM Cognition and Survey Methodology

CATI Computer assisted telephone interview

CIRF Cognitive Interviewing Reporting Framework

ESL Early School Leavers Indicator

ISCED International Standard Classification of Education

LFS Labour Force Survey

LLL Lifelong Learning indicator

NEET Not in Employment, Education and Training indicator

NSO National Statistics Office

PAPI Pen and paper interview

1. Research Objectives

The aim of this paper is to assess and improve the quality of the data collected for the non-formal education (COURATT) variable in the Labour Force Survey [LFS] and to address under-reporting issues. Cognitive interviewing was carried out to reduce the risk of collecting incomplete data for this variable. COURATT measures the respondents' attendance in courses, seminars, conferences or private lessons outside the regular education system within the last four weeks (Eurostat, 2017).

This paper follows the Cognitive Interviewing Reporting Framework [CIRF] proposed by Boeije and Willis (2013) on reporting cognitive testing studies.

2. Literature Review

Cognitive interviewing was developed as a formal method in the 1980s as part of a collaboration between cognitive psychologists and survey researchers known as Cognition and Survey Methodology [CASM] (Willis, 2006). This is a widely used method for evaluating survey questions and entails collecting additional verbal information about the respondents' thought processes (Priede & Farrall, 2011) and their survey responses (Gray, Blake, & Campanelli, 2014) while administering a survey questionnaire. In turn, this information is used to evaluate the quality of the response collected, to help determine whether the question is generating the information that its author intends (Gray, Blake, & Campanelli, 2014) and any difficulties that the participant has when answering the questionnaire (Priede & Farrall, 2011). This is also fundamental for judging whether survey results provide valid interpretations (Ryan, Gannon-Slater, & Culbertson, 2012).

Tourangeau (1984, as cited in Willis, 1999) developed a four-stage cognitive model to explore the phases of the question and answer process. These are question comprehension, retrieval from memory of relevant information, decision processes and response processes (Willis, 1999; Willis, 2006). The first component, comprehension, refers to whether the respondent understands the question in the same way as the researcher intended (Collins, 2003). Hence, attention is given to instructions and question wording, making sense of the question, determining what information is being asked, and making connections between key terms in the question and relevant concepts (Ryan, Gannon-Slater, & Culbertson, 2012). Once respondents understand the question, they usually have to retrieve information to answer it. The retrieval component processes include adopting remembering strategies, trying to activate recollections, recall memories, and arriving at conclusions that complete partial recollection (Ryan, Gannon-Slater, & Culbertson, 2012). Subsequently, respondents engage in judgment processes such as evaluating the importance and completeness of the recalled information, combining the information retrieved in order to answer the question, and estimating gaps to adjust for what is missing. After preparing a judgment, the respondent chooses an answer and communicates it (Ryan, Gannon-Slater, & Culbertson, 2012). Respondents may edit their answer prior to communicating to conform to notions of social desirability and self-presentation (Collins, 2003). Through cognitive testing techniques including verbal probing and/or thinking aloud, the interviewer should be able to understand these thought processes (Ryan, Gannon-Slater, & Culbertson, 2012).

3. Research Design

Although cognitive interviewing is mainly used before the main survey as part of the questionnaire development stage, Willis (2005) states that cognitive interviewing can be conducted at various points (Gray, Blake, & Campanelli, 2014). In this case, cognitive interviewing was used to test the non-formal education (COURATT) question and the respondents' answer. For the purpose of this study two questions were presented to the participants. The original question was broad and its clarifications were included in the adjacent questionnaire notes, which in most cases, were not read out by interviewers. The revised version was then divided into six detailed short questions with the primary intention of improving the probability that an event will be recalled and reported (Fowler, 1995) as well as to aid comprehension (Willis, 1999).

Original question

During the previous 4 weeks have you attended any courses, seminars, conferences or private lessons outside regular education?

Yes = 1
No = 2

Prior to testing, the question was re-designed in English following the examples proposed in the LFS Explanatory notes and then translated in Maltese (Eurostat, 2017).

Revised question

During the previous 4 weeks have you attended any of the following courses:

	Yes	No
(a) On job Training		
(b) Driving Lessons		
(c) Private Lessons		
(d) Seminars or conferences		
(e) Music, band clubs, singing, drama, dancing or sport lessons		
(f) Other courses (Including foreign language lessons, hobbies, art courses, health courses (e.g. natal courses), CANA Movement courses)		

In this study, a hybrid cognitive interviewing model was applied where probing techniques were used in conjunction to the think-aloud method (Ryan, Gannon-Slater, & Culbertson, 2012; Boeije & Willis, 2013). Initially, the respondents were asked the original question and answers were recorded by the interviewer. The respondents were instructed to think aloud while answering the question. In thinking aloud, the respondents are encouraged to vocalise their thought processes as they answer a survey item (Ryan, Gannon-Slater, & Culbertson, 2012). The interviewer took written notes of the processes the participant used in arriving at their answer to the question (Willis, 1999) and these were examined later on (Priede & Farrall, 2011). This approach helps the researcher to gain insight into the relevant cognitive processes that are taking place (Ryan, Gannon-Slater, & Culbertson, 2012).

Subsequently, the interviewer employed verbal probing to acquire more knowledge on the respondents' thought processes. This technique was used to target the various cognitive processes respondents engage into to finalise their answer, specifically, comprehension and recall (Ryan, Gannon-Slater, & Culbertson, 2012). For this study retrospective probing was used, where the subject was asked the probe after the question was answered (Boeije & Willis, 2013). This technique stimulates a more realistic environment; however, there is a chance for the respondents to forget what they were thinking as they answered the question (Willis, 1999). As only one question was being tested, retrospective probing was also applied to avoid interruption during the interview (Willis, 2006). Moreover, a combination of probing types was used, mainly scripted probes developed prior to the interview. However, spontaneous probes were also applied to follow-up on interesting issues that emerged through the course of the interview (Willis, 1999). The scripted probes used were oriented towards particular cognitive processes:

1. Comprehension: What do you understand by the word "courses"?
2. Recall: How did you remember that?
3. Confidence and judgement: How well do you remember this?

The respondent was then presented with the revised question. The interviewer took note of the responses and used general follow-up probes (How easy or difficult did you find this question to answer?) to draw comparison between both questions and explore relevant issues (Castillo-Díaz & Padilla, 2013).

4. Participant Selection

Cognitive interviews are not conducted for the purpose of generalizing or to statistically represent a larger population; hence there is no agreement about participant selection and adequate sample sizes (Ryan, Gannon-Slater, & Culbertson, 2012). For this study, 10 cognitive interviews were conducted as proposed by Willis (2006) and participants were recruited by word of mouth. The target population consisted of persons aged 15 years and over and respondents were selected through quota sampling to obtain a range of ages (3: 15-30, 3: 31-50, 4: 51+), sex (4: male, 6: female), and educational levels (3: low, 4: medium, 3: high) (Gray, Blake, & Campanelli, 2014). The purpose of this sample was to interview a variety of individuals rather than providing precise statistical estimation (Willis, 2006). Interviews were carried out in a meeting room within the National Statistics Office [NSO] which served as a "laboratory" to conduct interviews (Willis, 1999, p. 26).

5. Data Collection

In line with Willis (2006), the cognitive interviews were conducted by a single interviewer with experience in questionnaire design, cognitive processes and interviewing, and with experience in social science research concepts and having good inter-personal skills. Cognitive interviewing was conducted via both pen-and-paper interviewing [PAPI] and computer assisted telephone interviewing [CATI] modes of data collection to replicate survey conditions. In this way the key features that distinguish between different modes of interviewing were reproduced, specifically nonverbal communication in CATI (Gray, Blake, & Campanelli, 2014). The interviews were not recorded; however, the interviewer took field notes of the answer modalities, observations of non-verbal gestures and probes used during the course of the interview.

Since difficulties with the question were detected prior to cognitive testing, both the original and the revised version of the question were tested. In the 7th interview major problems related to the question and the necessary modifications were apparent and saturation was achieved. From the cognitive testing, results pertaining to the revised question were deemed sufficient to be sent into the field for actual data collection. Consequently, one round of cognitive testing was perceived to be sufficient.

6. Data Analysis

A qualitative analysis of the cognitive interviewing data was carried out to reveal difficulties respondents have with the survey context, understanding questions, retrieving information, and to provide answers in order to revise the question (Ryan, Gannon-Slater, & Culbertson, 2012). The researcher analysed the individuals' cognitive processing problems during individual interviews, consistencies, and inconsistencies in question response processes and patterns across interviews. Comparisons of subgroup response differences were also evaluated to explore potential bias, validity and reliability issues (Collins, 2003).

Since the outcome of the cognitive interviewing is qualitative, the researcher looked for dominant trends across interviews and grouped the acquired information into themes (Willis, 1999). The major problems deducted in this study were grouped in two themes:

Theme A: Difficulty in understanding the term "course"
Theme B: Issues with recalling participation in non-formal education in the past four weeks

After data was analysed, information regarding the cognitive testing analysis was provided and discussed with the other statisticians working on the LFS. Subsequently, the interpretation of findings and its effect on LFS data was presented to the senior management within the NSO.

7. Findings

The cognitive interviews described in this paper provided insight into respondents' thought processes while answering the question. This also contributed to the understanding of the difficulties within the quantitative data concerning the COURATT variable within the LFS questionnaire.

The largest issue identified with cognitive interviewing related to the comprehension phase. In fact, findings indicated lexical problems associated with respondents' understanding of the meaning and the use of the terms "courses, seminars, conferences or private lessons" and the context in which they are used in the questionnaire (Drennan, 2002). According to the LFS explanatory notes the scope of this question is to identify any of the following courses; taught courses on job-related skills; courses or seminars provided by an external enterprise and not the employer; study circles; training preparation for the labour market; courses, seminar or conference provided by trade unions or employers' associations; foreign language course; art courses; piano lessons; riding lessons; driving school; hobbies courses; courses for personal or social reasons and instructions or private lessons from a teacher or a tutor (Eurostat, 2018).

However, through cognitive testing, it became evident that when answering the question respondents were assuming that these courses should relate to their job or occupation, or to a formal seminar or conference. Consequently, the majority of the courses were not being considered when answering this question leading to under-reporting. The adopted multiple question format approach made it easier for the respondent to understand the scope of the question correctly and to provide reliable and valid data.

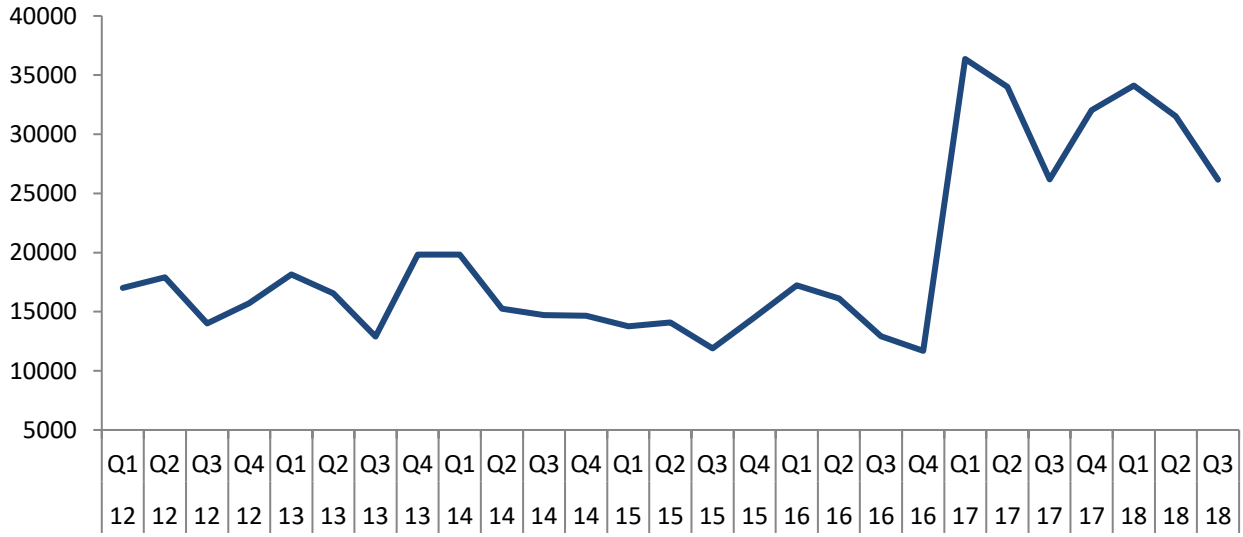
The second issue identified through cognitive testing related the retrieval stage, that is, how the respondents access the information in their mind, necessary for answering the question (Priede & Farrall, 2011). It was apparent that respondents had trouble retrieving relevant information needed to respond this question (Ryan, Gannon-Slater, & Culbertson, 2012). When uncertain, respondents were choosing the answer category "No", believing it to be the closest to what they knew, which resulted in an exclusion problem (Drennan, 2002). Hence, the multiple-question approach helped to ease the respondent's thinking about past experiences and served as a "carryover effect", where meaning was carried over from one question to the other (Ryan, Gannon-Slater, & Culbertson, 2012, p. 426). The revised question started with the most understood course type, that is, on-job training and proceeded with other courses to set it within a context and to aid respondents' recall.

8. Implications

The integration of cognitive theory with survey methodology provides a detailed understanding of the respondents' thought process in answering the questionnaire (Drennan, 2002). Following the cognitive testing implemented in the final quarter of 2016 for the COURATT variable, changes were introduced in the first quarter of 2017. As described in section 3, the question was divided into six short and detailed questions with the aim of covering all the courses defined by the LFS explanatory notes. A macro-level analysis on the quarterly time series data from quarter 1 2012 to quarter 3 2018 was carried out to assess the impact of the new question. The aim was to monitor and evaluate its effect and to determine the nature of change points in the first quarter of 2017 (Selukar, 2017). In our analyses, JDemetra+ was used to detect the structural break through the ARIMA models.

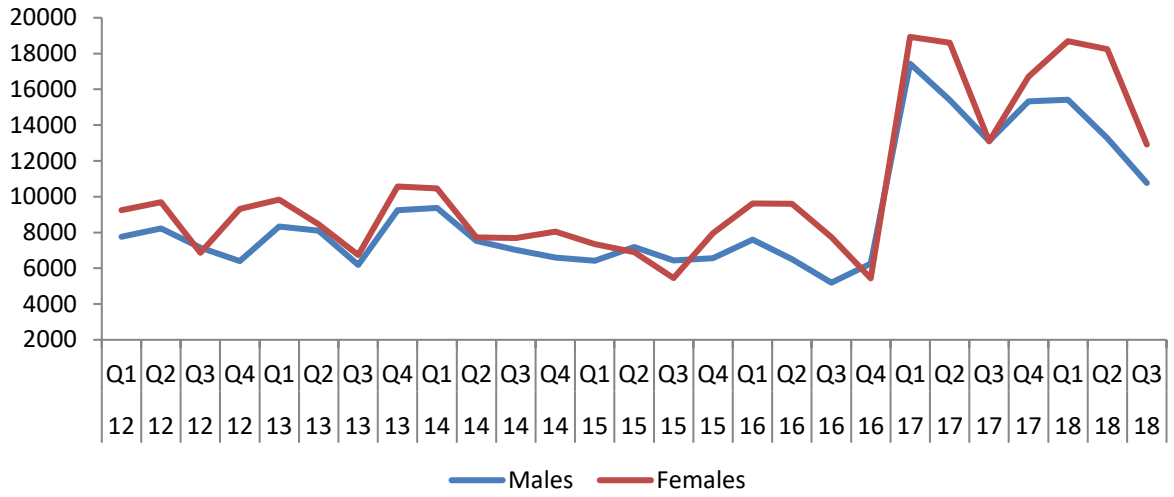
A visual inspection of the LFS non-formal education figures indicate a level shift of 24,674 between quarter 4 2016 and quarter 1 2017 (Chart 1). The statistically significant break was determined by the automatic outlier detection ($p < 0.001$).

Chart 1: The number of persons 15 years and over in non-formal education



Similar level shifts were recorded for both males and females attending non-formal education ($p < 0.001$). LFS data showed an increase of 11,176 and 13,948 persons respectively with the revision in the COURATT question (Chart 2).

Chart 2: The number of persons 15 years and over in non formal education by sex



Participation in non-formal education estimates were also compared across different socio-demographic variables including highest level of education attained, age, proxy interviewing and panel. For each socio-demographic variable, tests were carried out to compare the mean scores of quarterly data before and after the implementation of the new question. Since data was not normally distributed, the Kruskal-Wallis test was used to test the mean scores between three or more independent groups. Test results indicated that non-formal education figures were mostly significant for two out of three educational attainment categories. The low level of education category comprising of persons with ISCED 2 or less proved to be statistically significant ($p = 0.031$). In addition, the medium level of education category including persons with ISCED 3 and 4 also proved to be statistically significant ($p = 0.031$). This might imply that prior to the revision; respondents were associating the COURATT question only with job-related courses which are predominantly more common among persons with a high level of education.

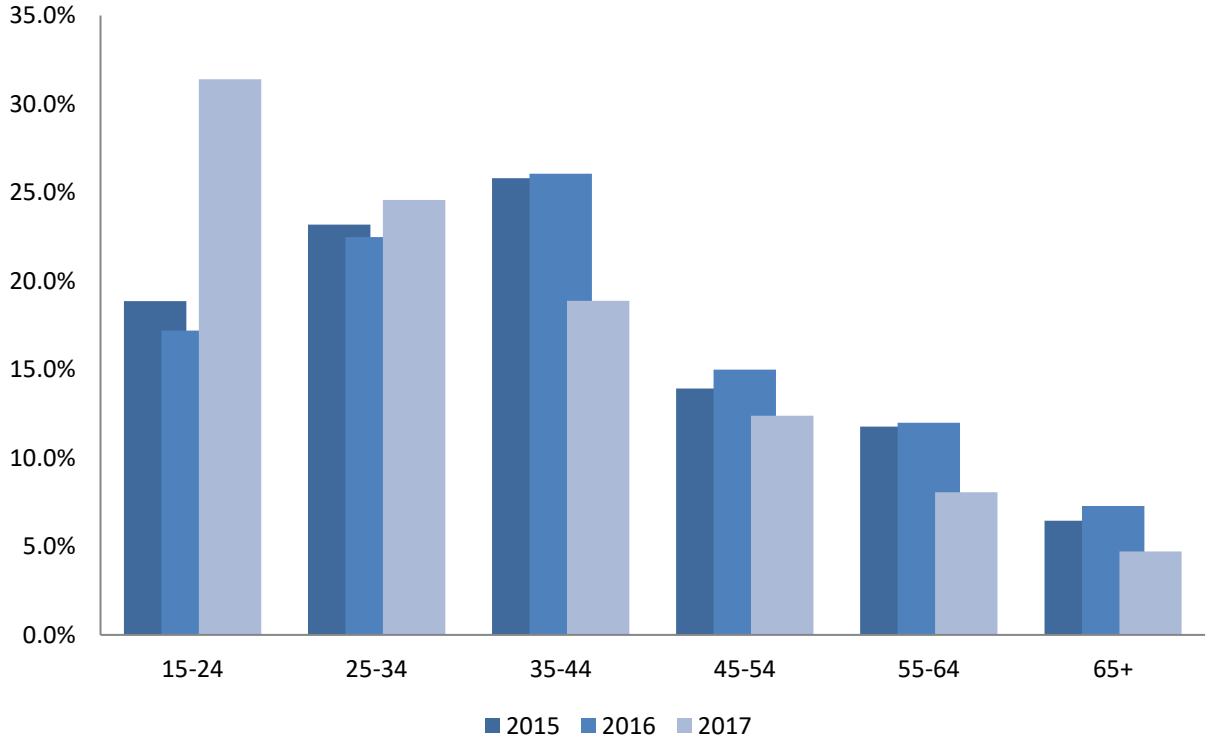
Furthermore, the Kruskal-Wallis test was used to check significant changes before and after the implementation of the new question across panels. Test results implied that the changes in the question affected the distribution of persons in non-formal education for the first and third panel ($p = 0.014$ and $p = 0.038$ respectively).

Participation in non-formal education levels were also checked by proxy interviewing. The Mann Whitney U-test, a non-parametric test analogue for independent samples t-test, was applied to compare the mean scores for direct interviewing only. Test results indicated a statistically significant difference in the underlying

distributions of participation in non-formal training for the direct interviews between Quarter 4 2016 and Quarter 1 2017 (p = 0.005).

Further analysis on this variable was carried out by age group. Results from the Kruskal-Wallis test showed statistically significant difference in non-formal education participation (before and after COURATT implementation) for the 15 to 24 and the 45 years and over age groups (p = 0.000). Meanwhile, test results showed no significant changes in the 25 to 44 age group who tend to participate more in job-related courses. This further substantiates the argument that prior to cognitive testing the COURATT variable was mostly covering job-related training since the majority of the participants understood this question in the context of work-related courses. Hence, there was no affect on this age group despite the revisions in the questionnaire (Chart 3).

Chart 3: The number of persons in non-formal education by age group: 2015 - 2017



An evaluation of the impact of change on education indicators related to the COURATT variable was carried out to monitor and determine the break point of the following indicators:

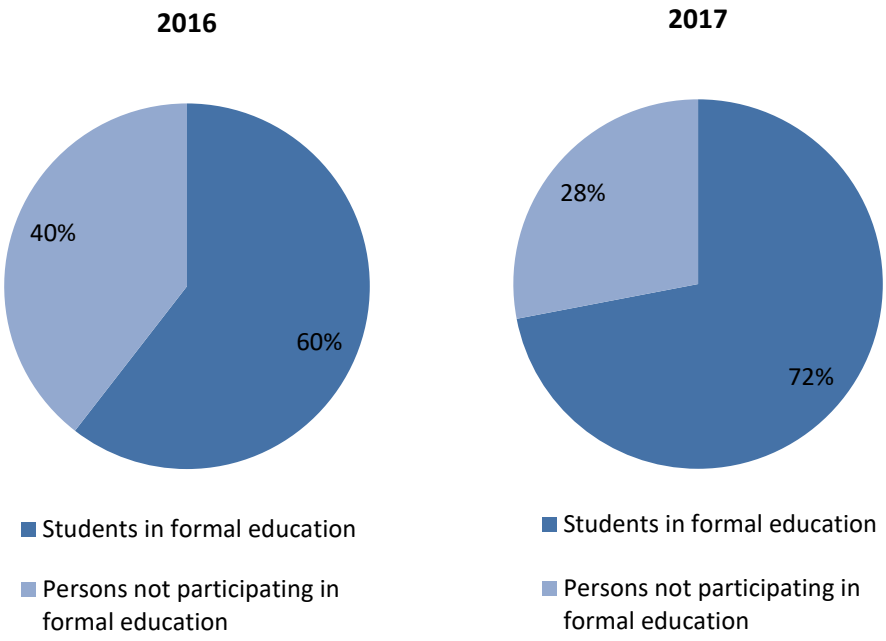
1. Early School Leavers indicator (ESL)
2. Not in employment, education and training indicator (NEET)
3. Lifelong learning indicator (LLL).

The automatic outlier detection based on iterative procedures carried out for the ESL and NEET indicators did not prove an unexpected shift in time series, hence a structural break was not reported. LFS estimates for 2016

and 2017 show that the majority of the 15 to 24 age group who were attending some form of non-formal education, were also enrolled in formal education (Chart 4).

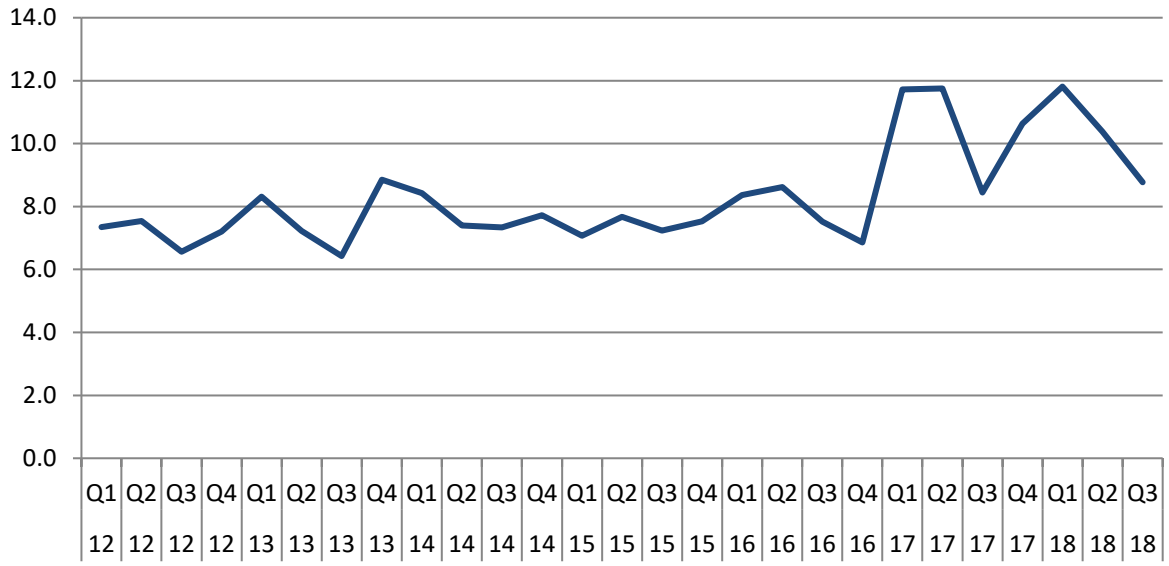
As a result, one may argue that, the revised COURATT question implemented in the first quarter of 2017 had no significant effect on the ESL indicator since the majority were already enrolled in formal education. This indicator covers persons aged 18 to 24 who achieved secondary education or less and who are not pursuing further education or training. Correspondingly, this modification did not impact the NEET indicator, that is, the number of persons aged 15 to 24 years who are not in employment, education or training as a percentage of total persons within this age category.

Chart 4: The percentage of persons aged 15 to 24 attending non-formal education by formal education participation: 2016 and 2017



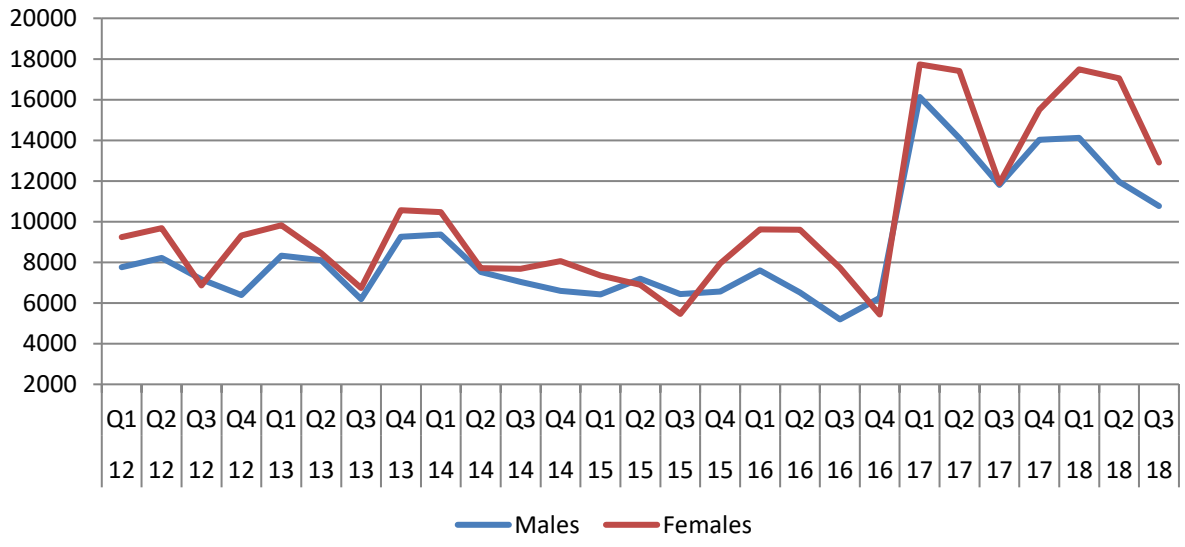
However, the implementation of the revised multiple question in Quarter 1 2017 resulted in a significant level shift break in series for the LLL indicator (Chart 5) as per the automatic outlier detection ($p < 0.001$). This indicator refers to the percentage of persons between 25 and 64 years participating in regular education, or in non-formal training such as courses, seminars and conferences. In the first quarter of 2017, a level shift of 4.9 percentage points was reported when compared to the previous quarter in 2016. This suggests that cognitive testing aided the understanding of the question (Willis, Cognitive interviewing: A “how to” guide, 1999) and helped the respondent to remember his or her participation in non-formal training (Fowler, 1995).

Chart 5: The percentage of persons in LLL



The level shift in the LLL indicator proved to be significant for both sexes ($p < 0.001$). The estimated change between both data points were 4.0 and 5.8 percentage points for males and females respectively (Chart 6).

Chart 6: The percentage of persons in LLL by sex



10. Strengths, limitations and recommendations

The strongest feature of this study is the methodology used, as cognitive testing improves our understanding of measurement error in quantitative surveys (Collins, 2003) and enhances the validity of survey instruments (Hofmeyer, Sheingold, & Taylor, 2015). A further strength of this research is that this method proved to be

reliable for research to assess whether the study objectives are achieved. Through the process of cognitive interviewing, question completion can be viewed from the perspective of the respondent rather than the researcher. This allows a better insight into the participant's problems which may have not been anticipated prior to the distribution of the questionnaire. It also ensures data compatibility where respondents interpret the questions in the same way (Drennan, 2002). Consequently, cognitive testing also helped to develop better questions and survey instruments and to pre-test questionnaires (Collins, 2003).

Meanwhile, cognitive interviewing has been criticized for creating a false environment and for adding a "cognitive load" to respondents resulting in "artificiality" of the overall process (Drennan, 2002, pp. 61-62). The most challenging aspect of this study was that respondents were forced to think about the question in a different manner and for a longer period than usual, which, in itself may have changed the answer and attitude to the question (Priede & Farrall, 2011).

Another weakness of this study concerns its qualitative nature as it is subjectively based on the researcher's interpretations (Drennan, 2002) and mainly relies on the respondents' verbal reports of problems (Collins, 2003). Finally, because cognitive methods are fairly new, they are still relatively non-standardised hence results are not always reliable, particularly those based on cognitive interviews (Collins, 2003).

Based on findings generated from this study, the researchers suggest the implementation of cognitive testing at the pre-testing stage in other social surveys. Cognitive interviewing provided significant information on the respondents' thought processes and questionnaire design. A further recommendation is the application of cognitive testing in ad-hoc questionnaires and in the design of the LFS questionnaire for the parallel run as part of the new LFS regulation.

10. Conclusion

This cognitive testing exercise was carried out to evaluate non-formal education (COURATT) estimates in the LFS. The aim was to determine the respondents' thought processes in answering this question and to improve under-reporting. Findings from this study contributed to the enhancement of the question under-study and to comprehend the difficulties within the quantitative data concerning this variable. Due to the changes implemented in the questionnaire, a break-in-series was detected in the non-formal education estimates and a level shift was recorded in the LLL indicator.

References

Boeije, H., & Willis, G. (2013). The cognitive interviewing reporting framework (CIRF): towards the harmonization of cognitive testing reports. *European Journal of Research Methods for the Behavioral and Social Sciences Vol. 9 (3)*, 87-95.

Castillo-Díaz, M., & Padilla, J.-L. (2013). How cognitive interviewing can provide validity evidence of the response processes to scale items. *Social Indicators Research Vol.114 (3)*, 963-975.

Collins, D. (2003). Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research Vol. 12*, 229-238.

Drennan, J. (2002). Cognitive interviewing: verbal data in the design and pretesting of questionnaires. *Methodological Issues In Nursing Research Vol. 42 (1)*, 57-63.

Eurostat. (2017). *EU Labour Force Survey Explanatory Notes.* Luxembourg: Eurostat.

Eurostat. (2018). *Handbook on Seasonal Adjustment.* Luxembourg: Publications Office of the European Union.

Fowler, F. (1995). *Improving Survey Questions: Design and Evaluation.* Thousand Oaks: California.

Gray, M., Blake, M., & Campanelli, P. (2014). The use of cognitive interviewing methods to evaluate mode effects in survey questions. *Field Methods Vol. 26(2)*, 156-171.

Hofmeyer, A., Sheingold, B. H., & Taylor, R. (2015). Do you understand what I mean? How cognitive interviewing can strengthen valid, reliable study instruments and dissemination products. *Journal of International Education Research Vol.11 (4)*, 261-268.

Priede, C., & Farrall, S. (2011). Comparing results from different styles of cognitive interviewing: 'verbal probing' vs. 'thinking aloud'. *International Journal of Social Research Methodology Vol. 14, No. 4*, 271-287.

Ryan, K., Gannon-Slater, N., & Culbertson, M. J. (2012). Improving survey methods with cognitive interviews in small- and medium-scale evaluations. *American Journal of Evaluation 33(3)*, 414-430.

Selukar, R. (2017). *Detecting and adjusting structural breaks in time series and panel data using the SSM procedure.* Cary, North Carolina: SAS Institute Inc.

Willis, G. (1999). Cognitive interviewing: A “how to” guide. *Reducing Survey Error through Research on the Cognitive and Decision Processes in Surveys* (pp. 1-41). Alexandria, Virginia: Research Triangle Institute.

Willis, G. (2006). Cognitive interviewing as a tool for improving the informed consent process. *Journal of Empirical Research on Human Research Ethics*, 9-24.