Organisation for Economic Co-operation and Development

COM/DELSA/EDU/PIAAC(2018)6

For Official Use **English - Or. English**

27 March 2018

Directorate for Education and Skills
Directorate for Employment, Labour and Social Affairs
Board of Participating Countries for the Programme for the International Assessment of Adults Competencies

THE USE OF TABLETS FOR COLLECTING DATA IN THE 2ND CYCLE OF PIAAC

20th meeting of the PIAAC Board of Participating Countries

16-17 April 2018, OECD Headquarters, Paris.

Members are asked to:

NOTE and DISCUSS the content of the paper.

APPROVE the proposal concerning the use of delivery devices in the 2nd cycle of PIAAC.

Andreas Schleicher [Andreas.Schleicher@oecd.org; Tel: +33 (0) 1 45 24 93 66]
William Thorn [William.Thorn@oecd.org; Tel: +33 (0) 1 45 24 78 04]

JT03429057

Table of contents

1. Introduction **3**
2. The tablet-only design **4**
3. The tablets: Technical specifications, user interface and item types **6**
4. The feasibility of a tablet-only design **7**
5. The option of paper administration **9**

1. Introduction

1. The Terms of Reference for the 2nd cycle of PIAAC explicitly invited bidders to make "as much use as possible of appropriate digital technologies", in order to "increase the efficiency and quality of the collected data", while at the same time paying attention to the needs of respondents which might lack familiarity with digital devices and applications.

2. Greater reliance on digital devices during the data collection phase was seen as instrumental to some important objectives of PIAAC, such as:

- Innovation, in particular, to continue to be at the leading edge in terms of the methodology and delivery of household surveys and large-scale assessment;
- Increase the efficiency of data collection, reducing the potential for mistakes and improving the monitoring of survey administration tasks;
- Improve the data quality and the measurement of the underlying cognitive constructs, by increasing the share of respondents taking the full assessment on the same device or in the same mode; and
- Increase the information collected across all participants, such as timing and process data that can be used both to support survey management and quality assurance and to provide information reflecting on the strategies and behaviours of respondents.

3. The selected international contractor proposed two options for the assessment design: tablet-only administration and laptop plus paper administration (essentially identical to that adopted in the 1st cycle of PIAAC). In discussing, the proposal at the 19th meeting of the BPC, a number of countries expressed some doubts and concerns on a tablet-only administration and asked that the possibility of tablet-based administration complemented with a paper-based assessment for adults unable or unwilling to rake the assessment on the tablet be considered.

4. This paper provides more details on the use of tablets for test administration (highlighting the pros and cons of this option). It recommends that the BPC agree to:

- Adopt a tablet-based delivery in the Field Trial.
- Leave open the possibility of offering the paper-and-pencil assessment of literacy and numeracy that was used in the 1st cycle for respondents who lack familiarity with digital devices in the Main Study only if the Field Trial reveals major problems associated with the tablet delivery.

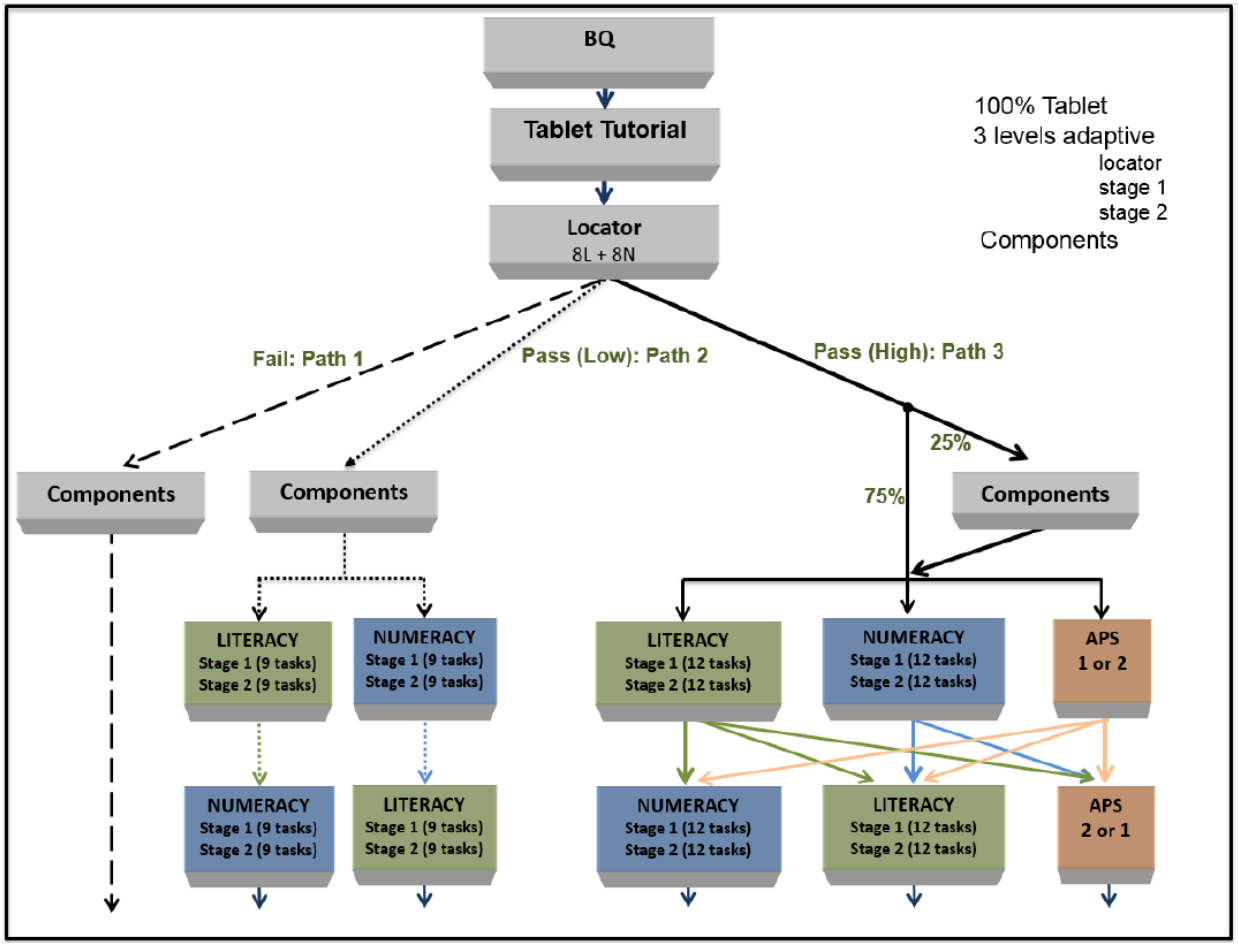
5. Members are asked to:

- NOTE and DISCUSS the content of the paper.
- APPROVE the proposal concerning the use of delivery devices in the 2nd cycle of PIAAC.

2. The tablet-only design

6. A proposed tablet-only Main Study design is presented in Figure 2.1. It is important to note that this design fulfils the set of design criteria presented in the Terms of Reference and is subject to alteration following the meetings of the various expert groups. All adults agreeing to participate in the survey will be asked to respond to a set of questions on familiarity with electronic devices as part of the Background Questionnaire (BQ). After completing the BQ respondents will receive a tutorial providing them with an opportunity to practice the range of item types used in the PIAAC direct assessment. This will be followed by a locator test that will include eight literacy and eight numeracy items.

Figure 2.1. The tablet-only design



Source: Proposal submitted by ETS in response to the PIAAC Call for Tenders RFQ 198.

7. Depending on the information collected via the BQ and locator test, respondents will be directed to one of three paths. The information obtained during the Field Trial will be used to calibrate the locator test in order to decide the cut-off to be used to route respondents to the different paths.

8. Respondents who fail the locator will only take the assessment of reading (and, if developed by the expert group, numeracy) components. Respondents who pass the locator with a low score will take the assessment of components, followed by an adaptive assessment of literacy and numeracy. Respondents who pass the locator with a high score will take a multi-stage adaptive assessment of literacy, numeracy and adaptive problem solving (APS). It is worth noting here that it will be up to the expert group to determine whether or not APS will be adaptive and that adaptive testing will only be used in the Main Study. Before receiving the assessment of literacy, numeracy and APS approximately one quarter of respondents in this last path will be randomly selected to also take the components assessment so that performance on the components can be generalised to the overall population.

9. This design is seen as best meeting the stated objectives of PIAAC. In particular, a tablet-only design will have the following advantages:

1. It will improve efficiency and improve the quality of the data, by allowing for automatic scoring of all domains in the assessment. Automated scoring will also contribute to a reduction in national costs, as there will be no need to score items manually, and therefore no need to hire and train scorers or for individuals to manually enter data.
2. It will improve the measurement properties of the assessments, by removing any item-by-country interactions due to differences in the proportion of the population using different modes of delivery in different countries.
3. It will increase the measurement of the proficiency scales by having all adults responding to the same set of tasks, which results from the fact that all newly developed assessment tasks will be electronic only
4. Delivering the component assessment on a digital device will also provide richer information, e.g. through timing data, which is important to more accurately assess the fluency of respondents.
5. It will take full advantage of multi-stage adaptive testing for all respondents.
6. It will still maintain the measurement of trends in literacy and numeracy back to the 1st cycle of PIAAC cycle and previous adult surveys (IALS, ALL).
7. It will maximise the information on the covariance structure among the various measures of proficiency.
8. It will maximise the amount of paradata (log-files) that can be obtained from the study.

10. Two potential disadvantages of adopting a tablet-only design can be identified:

1. The cost of the devices, which should however be weighed against the savings associated with the automatic scoring of all items (e.g. no need to hire staff to input the scored data manually), as well as against the cost of traditional laptops that were used in the 1st cycle.
2. The risk that some respondents, not at ease with the use of digital devices, will refuse to take the assessment on the tablet or that their performance will be somehow negatively affected. However, this risk will be evaluated as part of the FT design.

11. The remainder of the paper will provide more details on the characteristics of the tablets, the user interface, and on the steps that the Contractor plans to undertake to minimise the risks and possible disadvantages associated with the use of tablets.

3. The tablets: Technical specifications, user interface and item types

12. The tablets will have to meet the following technical specifications:

- 11.5-inches screen or larger;
- Operating under Windows;
- Support input through a digital stylus;
- Support a keyboard, to be used by the interviewer for administering the BQ.

13. It will be possible to use hybrid devices, i.e. laptops with a detachable touch screen that can be used as a tablet once it is detached. These devices are becoming increasingly common, and they are now typically more rugged and more easily available than tablets. Additionally, "2 in 1 laptops", where the keyboard can be rotated around under the screen, are a possibility. Further usability testing of these devices must be done. More information on these types of hybrid devices can be found in [COM/DELSA/EDU/PIAAC(2017)4].

14. Tablet-based delivery will support a range of item types, including interactive items and simulations. All response modes used in the 1st cycle will still be available. The selected trend items for literacy and numeracy can be migrated to the tablet, and their performance, along with that of all newly developed items, will be evaluated in the FT.

15. Respondents will have to perform two simple actions when interacting with the tablet: tapping and dragging.

16. Both tapping and dragging can be performed using either a finger or the stylus. Palm-rejection technology ensures that no input is inadvertently recorded once the stylus is sufficiently close to the screen (i.e. only the input of the stylus is recorded). Palm-rejection technology minimises also mistakes and greatly improve the user experience, as respondents can put their hand or palm on the screen and record a response as they would do using a normal sheet of paper.

17. Tapping will be used to select information already present on the screen. Item types requiring tapping include:

- Multiple-choice (both single and multiple selection)
- Numeric entry (selecting numbers or symbols on a keypad)
- Hot spots (selecting areas within a text, a graph or an image)
- Radio buttons (selecting one or more choices or values)
- Selecting links on a web page or an email.

18. Dragging will involve moving objects from one location on the screen to another. Item types requiring dragging include:

- Drag and drop, or moving images or text boxes on the screen. This can be used to show sequences, association and matching, categorization, or fill in the blank.
- Slider, or moving a marker to indicate a numerical value between a lower and an upper bound.

19. For trend literacy items that use the highlighting response mode, respondents will simply be able to use the stylus to underline text, just as they would do on paper. A demonstration of the functionality of the tablets will be undertaken by staff from the consortium as part of this agenda item and consortium staff will be available to provide an hands-on demonstration of the functionalities of the tablets, showing some sample item types (that will be designed to resemble as much as possible what final PIAAC items are likely to look like), and answering more detailed questions on the technical specifications that will be required of the tablets during the meeting.

4. The feasibility of a tablet-only design

20. In the 1st cycle of PIAAC, approximately three-quarters of respondents took the assessment on a computer: 10% of respondents were directed to the paper-based assessment because they reported no prior computer experience, another 10% chose to opt-out of the computer assessment (without even taking the computer-based assessment core), and 4% failed the computer-based assessment core.

21. This highlights how only 4% of respondents actually demonstrated insufficient ability to take the assessment on the computer. The design of the survey meant that not all respondents took the ICT core. Based on their responses to the ICT use questions, it is likely that some respondents who took the paper-based assessment would have been perfectly capable of taking the computer-based assessment.

22. More importantly, it is expected that the number of respondents not familiar with digital devices will have substantially decreased over the ten years that separate the 1st and 2nd cycle of PIAAC, for two reasons. First, digital devices (and in particular smartphones and tablets) have become increasingly popular in recent years across the entire population. Second, experience and familiarity with digital devices is highly correlated with year of birth. The cohort aged 55-65 years in the 1st cycle of PIAAC will, for example, no longer be part of the PIAAC target population in the 2nd cycle. It can be reasonably expected that age differences in familiarity with digital devices will be smaller in the 2nd cycle.

23. The choice of tablets as the delivery device (as opposed to the laptops that were used in the 1st cycle) is also intended to make the computer-based assessment accessible to as large a share of the population as possible. The share of adults that is familiar with touch-screen devices such as smartphones or tablets is arguably much higher than the share of adults that is familiar with using a traditional laptop. One of the reasons behind the success of smartphones is the ease of use that the touch-screen technology brings compared to devices that need to be operated using a mouse and a keyboard.

24. When tablets are paired with use of a digital stylus, the user experience is similar to the interaction with a normal sheet of paper. The user experience will be further improved in two ways. First, the choice of a minimum screen size of 11.5 inches means that the size of the screen will be very close to that of an A4 sheet of paper (whose diagonal is around 14 inches). Second, in order to ensure automatic scoring of all items, no text entry items will be used in PIAAC meaning that a keyboard will not be needed to take the assessment. This is important not only because many respondents may not be familiar with typing on a physical keyboard, but also because digital pop-up keyboards that are typically used to type text in smartphones or tablets reduce the amount of material that can be displayed on the screen, and are often less efficient and more difficult to use than traditional physical keyboards.

25. Numerical entry will, on the contrary, be supported by a pop-up calculator. However, a pop-up calculator takes much less screen space than a keyboard, and tapping on a calculator to insert a number or perform a basic arithmetical operation is not very different than operating a normal calculator, and is certainly much easier than typing text on a digital pop-up keyboard.

26. In addition the consortium has the experience of using tablets in PISA-D with out of school populations. Thus far several thousand young adults who are out of school have taken the PISA-D assessment with no reported problems. It is worth noting here that a subset of the PISA-D reading and mathematics items were taken from earlier adult assessments.

27. As part of the development of the user interface, the Consortium will conduct some small-scale usability testing. This usability testing will use a number of cognitive items that reflect a range of response modes, and will be useful to inform the development of test items and of the user interface, and in particular to identify essential information and practice that should be included in the tablet tutorial.

28. As noted above, a tutorial in the use of the tablet response modes is an integral component of the PIAAC design, and will provide respondents with an opportunity to practise and familiarise with the tablet before undertaking the PIAAC assessment.

29. The Field Trial will be the major source of information regarding user acceptance of tablet-delivery, most importantly of the extent to which the tablet can be used by all members of the target population. In order to get empirical information regarding the usability of the tablet interface it is important that the capacity of the entire target population to use the tablets is tested. For this reason, it is recommended that tablets only are used for data collection in the Field Trial. Key indicators will be refusal rates and item non-response rates.

30. The Field Trial data will also be analysed in order to assess the existence of any effects of the change in delivery mode from the 1st and 2nd cycles. The characteristics of trend items will be compared between the 1st cycle (where a laptop was used) and the FT of the 2nd cycle (where the tablet will be used). The absence of mode effects between the paper-and-pencil and the laptop has already been established in the 1st cycle and can be taken for granted, especially in light of the increased diffusion of ICT and digital devices that has taken place in the last years.

5. The option of paper administration

31. In the event that the Field Trial provides evidence of problems related to the tablet administration (e.g. high refusal or low item response rates from certain groups of respondents), it will be possible, as a fall-back option, to provide a paper-based assessment of literacy and numeracy for respondents who lack sufficient ICT skills to take the assessment on the tablet.

32. Opting for a mixed tablet and paper-based administration will have a number of additional drawbacks:

1. It will prevent the full benefits of a tablet-only administration (highlighted in Section 2) to be realised.
2. It will not support the assessment of Adaptive Problem Solving. This will lead to some of the problems that were encountered with the assessment of Problem Solving in Technology-Rich Environments in the 1st cycle being repeated, in particular the problem of comparing results across countries because of variation in the proportions of the population taking the assessment.
3. As this option would entail use of the paper-based literacy and numeracy assessments from the 1st cycle without further modification (as no new items will be developed for the paper-based instruments), it will mean that a proportion of the population will be assessed using an assessment that does not fully reflect the frameworks developed for the 2nd cycle particularly in numeracy where considerable revision of the assessment framework is expected.

33. In the event that a paper-based option is considered essential in the Main Study, it will be important that the study design minimises the number of respondents who would be directed to the paper-based option. For instance, contrary to the approach in the 1st cycle, all participants, irrespective of self-reported experience with digital devices, could be required to take the tablet tutorial and attempt the locator items, with only those respondents who scored poorly on the locator being given the option to take the paper-based assessment. Alternatively, the option of opting for the paper-based assessment could be open to only a pre-specified subgroup of the target population, whose characteristics can be defined on the basis of the results of the Field Trial.

34. It is also important to remember that some modules of the Background Questionnaire will be developed for self-administration (the module on social and emotional skills, for example). Self-administration of the BQ will give respondents a further possibility to practice with using the tablet before undertaking the cognitive assessment. Using self-administration during the Field Trial will provide further evidence about the ability of respondents to use the tablet for entering information.

35. Irrespective of the decisions that the BPC will take regarding the delivery mode for the cognitive assessment, the modules of the BQ that will be designated for self-administration will be undertaken on the tablet. Having some modules of the BQ filled in on a tablet or computer and some modules completed on paper will be highly impractical

and will dramatically increase the probability of mistakes in the data entry process. Moreover, the BQ is much less cognitively demanding, and respondents that have low familiarity with a digital device could always ask for the help of the interviewer who has much more freedom to intervene during the administration of the BQ than during the administration of the cognitive assessment.