How much time do adults spend on the PIAAC assessment and why does it matter?

Adult Skills in Focus #10

OECD
BETTER POLICIES FOR BETTER LIVES

- Administering assessments like PIAAC on a computer (rather than on paper) allows to record information on how respondents interact with the testing platform, and in particular how much time they spend on the various assessment tasks and items.
- This additional information, recorded in log-files, can be an important complement to traditional information on how respondents perform on the test. It can be used to characterise the cognitive strategies followed by respondents and help them improve. It can also be used to construct indicators of engagement (to proxy the effort exerted by respondents), which is particularly important to interpret performance in low-stakes assessments.
- There are large cross-country differences in the amount of time respondents spend on the assessment. Time spent tends to increase with the age and education level of respondents. Similar differences emerge in the incidence of disengagement. Disengagement is more prevalent in the second half of the assessment.

Computer-based administration of large-scale assessments, made possible by technological developments and the increasing familiarity with computers and digital devices, makes survey administration more efficient and reduces the risk of human error. More importantly, it allows collecting a richer set of information on test-takers: log-files can store not only the final answer to a specific assessment item, but also all interactions between the respondent and the testing platform. As all events recorded are associated with a timestamp, it is possible to compute the amount of time elapsed between these events.

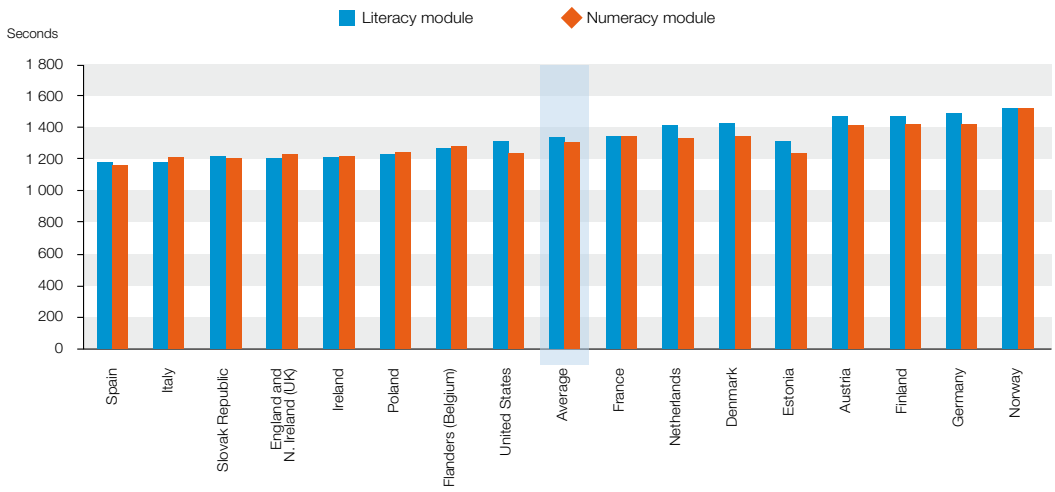
The interpretation of these data is complicated by the fact that many of the cognitive and non-cognitive processes respondents follow do not require an interaction with the testing platform. This is partly because most assessment items were not designed to allow the elicitation and the recording of specific actions or processes through the testing platform. Future assessments will leverage on the possibilities offered by digital testing interfaces and will deliver more informative data.

Still, relatively simple indicators such as the amount of time respondents spent on assessment items can be already fruitfully analysed and can provide valuable information that complement data on overall performance on the test, at both the individual and the country level.

There are large differences in the time respondents spend on the assessment

Respondents in Austria, Finland, Germany and Norway took the longest time to complete the literacy and numeracy assessments (about 50 minutes on average). In Ireland, Italy, Slovakia, Spain, England and Northern Ireland (United Kingdom), respondents spent about 40 minutes on average (Figure 1).

Figure 1 / **Total time on task in literacy and numeracy**



Note: The sample includes only participants to the computer based assessment who were assigned to the literacy and numeracy modules.

Source: OECD (2017), Programme for the International Assessment of Adult Competencies (PIAAC), log files. GESIS Data Archive, Cologne. ZA6712 Data file Version 2.0.0, doi:10.4232/1.12955

© OECD 2019

Part of these differences might be due to the fact that different respondents were typically assigned different items: the assessment was in fact adaptive, meaning that more able respondents were more likely to face more difficult items, and less able respondents more likely to face easier items.

One way to account for the adaptive nature of the assessment consists in computing, for each individual and for each item to which he/she was assigned, a position (expressed in percentile) in the overall distribution of time on task for each item, and then to average these percentiles across items. As most individual differences cancel away at the country level, this has very little effects on country rankings.

More interestingly, differences emerge in time on tasks according to individual socio-demographic characteristics. In particular:

- Higher-educated respondents spent more time on the assessment;
- Older respondents spent more time on the assessment;
- Respondents who declared to be more familiar with ICT spent less time on the assessment;
- Women spent less time than men, but gender differences were generally very small.

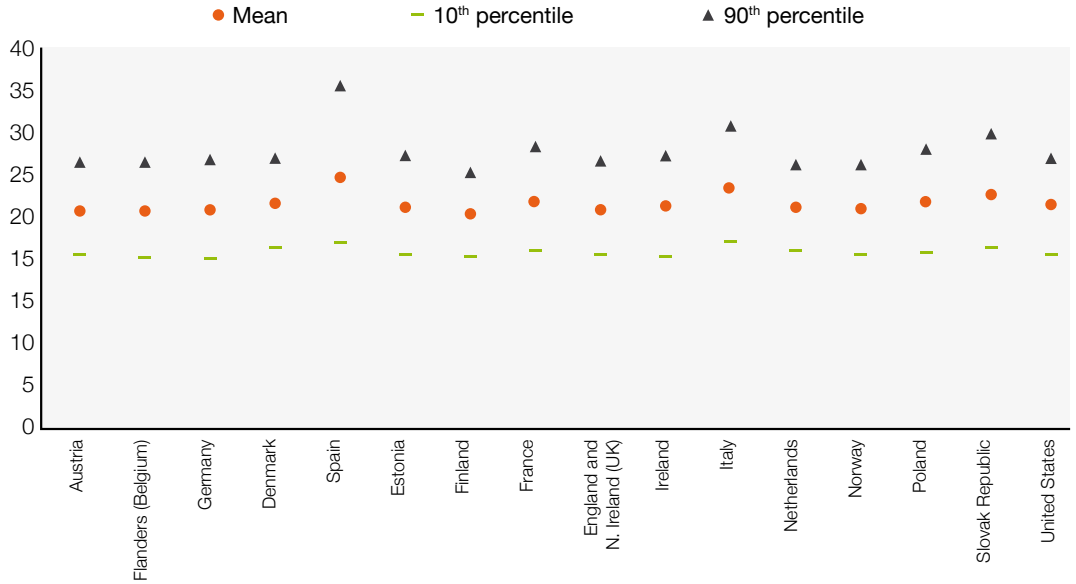
How do respondents allocate time to different items?

In order to analyse the time-allocation strategies followed by respondents in the course of the assessment it is useful to exploit the fact that, in the data, the same respondent is observed facing multiple items. This allows to hold constant all individual characteristics and comparing the behaviour of the same individual across different items.

A first question to address is whether respondents differ in the strategy they choose to allocate time between items. One way to answer this question is to compute, for each respondent, his/her position in the distribution of time on task for each item. Each respondent will then be characterised by a distribution of percentile ranks.

Respondents who always devote a similar amount of time (relative to other respondents assigned to the same items), will have a very compressed distribution of percentile ranks, while the distribution for respondents who spend an unusually large amount of time on some items and an unusually small amount of time on other items will be more dispersed. The average standard deviation of these respondent-specific distributions of percentile ranks is around 20 percentile points, pointing to a relatively low degree of individual consistency: it is not rare that respondents are relatively fast on some items and relatively slow on other items. On the other hand, there is very little cross-country variation in this indicator, as shown in Figure 2.

Figure 2 / **Individual distribution of time on task**



Note: The figure shows moments of the within-country distribution of individual standard deviations in percentile ranks. The sample includes only participants to the computer based assessment who were assigned to the literacy and numeracy modules.

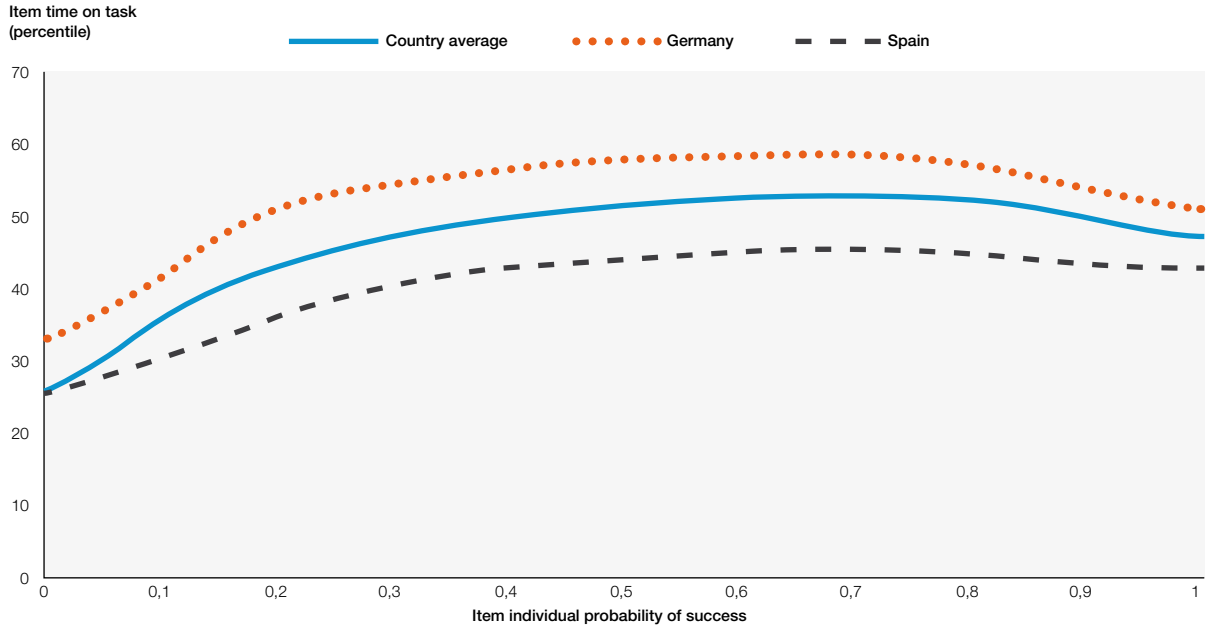
Source: OECD (2017), Programme for the International Assessment of Adult Competencies (PIAAC), log files. GESIS Data Archive, Cologne. ZA6712 Data file Version 2.0.0, doi:10.4232/1.12955

© OECD 2019

It is also possible to investigate how the time devoted to an item varies with the characteristics of the item itself, for instance item difficulty. An individual who values his or her time is not expected to devote a lot of time to questions that are too difficult, and which he/she is therefore very unlikely to be able to answer correctly. PIAAC allows calculating an individual specific ex-ante probability of success by combining information on respondents' proficiency and the "objective", pre-specified difficulty level of an item.

Figure 3 shows that, as items becomes excessively difficult with respect to the respondents' proficiency level, respondents devote less time to it (relative to more able respondents who faced the same item). Time-on-task percentiles also tend to decrease, although to a lesser extent, when items are very easy. The decline in time on task is lower at the top than at the bottom end of the probability of success distribution, because respondents are more likely to skip difficult items (therefore devoting very little time to them). Easy items, on the other hand, necessarily take some time to answer correctly.

Figure 3 / **Time on task and relative probability of success**



Note: The sample includes only participants to the computer based assessment who were assigned to the literacy and numeracy modules.
Source: OECD (2017), Programme for the International Assessment of Adult Competencies (PIAAC), log files. GESIS Data Archive, Cologne. ZA6712 Data file Version 2.0.0, doi:10.4232/1.12955

The possibility to observe the same individual answering multiple items also allows estimating the causal effect of spending more time on an item on the individual performance in the test, measured by the (ex-post) probability of answering the item correctly. On average, spending one additional second on an item increases the probability of giving a correct answer by 0.6 percentage points.

How much do respondents take the assessment seriously?

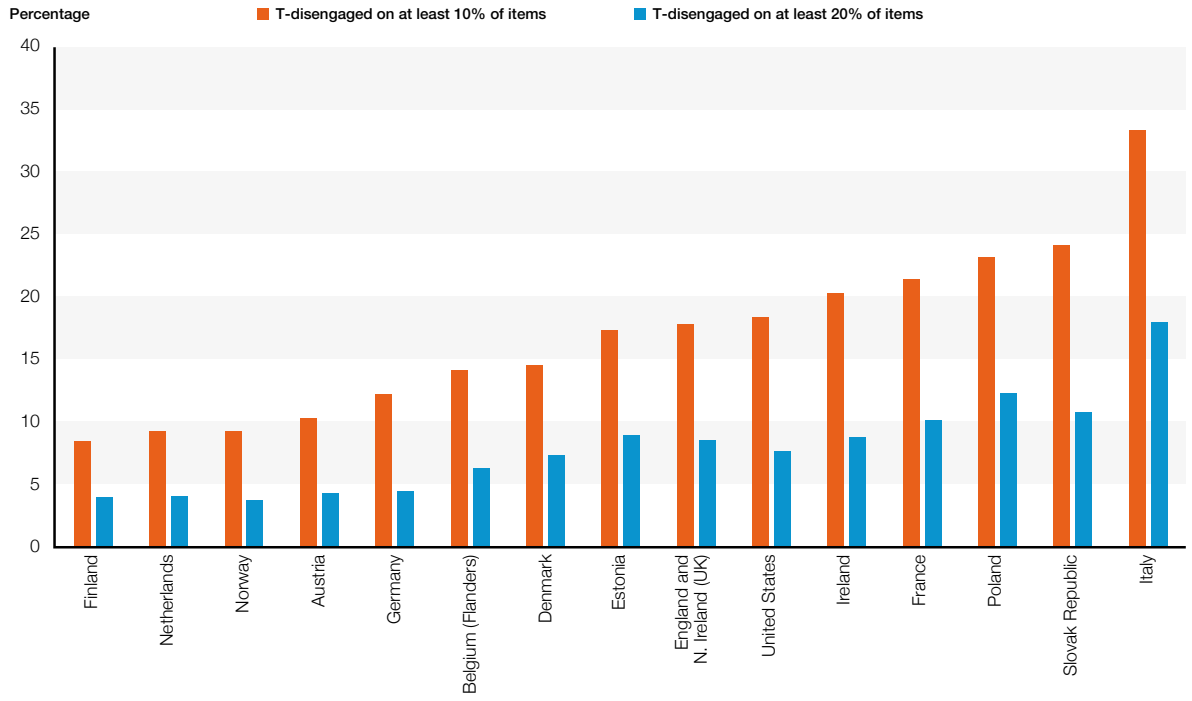
PIAAC is a low-stakes assessment. Unlike exams or competitions, performance in PIAAC has no consequences for individual respondents, who have then no external incentives to exert high levels of effort. If some respondents do not take the assessment seriously and do not perform at their best, a source of undesirable variation (or noise) is introduced in the final estimates of individual- and population-level proficiency. If different subgroups of respondents within and between countries display different levels of disengagement, the validity of comparisons between groups of respondents and between countries might be compromised.

The relationship between disengagement and performance is a complex question. First, disengagement in PIAAC can only be measured with indicators that only partially capture the spectrum of disengagement. Second, disengagement and low performance are linked in a complex relationship that cannot be easily disentangled. Third, PIAAC proficiency scores already partially account for disengagement by ignoring items on which respondents spent less than five seconds without giving an answer.

© OECD 2019

One way to classify respondents as disengaged with an item is to compute, for each item, the minimum amount of time that a respondent should spend on it in order to at least attempt to understand and solve the task (Goldhammer et al., 2016). Respondents spending less time than this threshold are labelled T-disengaged with that particular item. Figure 4 shows the proportions of the population who T-disengaged on at least 10% and at least 20% of items. Disengagement concerns respondents in all countries, but to a varying extent. Disengagement is much less frequent in northern European countries, such as Finland, Norway or the Netherlands. In these countries, about 8% of the sample disengage on at least 4 items out of 40. This proportion approaches 35% in Italy. The same differences between countries emerge when looking at more severe cases of disengagement, in which respondents disengage on at least 20% of items. The proportion drops below 5% in Finland, Norway and Netherlands, but it remains above 15% in Italy.

Figure 4 / **Disengagement rates across countries**



Source: OECD (2017), Programme for the International Assessment of Adult Competencies (PIAAC), log files. GESIS Data Archive, Cologne. ZA6712 Data file Version 2.0.0, doi:10.4232/1.12955

Respondents are more likely to be disengaged with items that appear in the second half of the assessment rather than in the first. This is consistent with the fact that respondents tend to spend less time on items positioned in the second module. This could partly be due to the fact that respondents become more familiar with the user interface and thus become quicker to solve the items. However, the increase in disengagement rates provides also evidence that the motivation of respondents, or their ability to endure effort for a prolonged amount of time, decreases with the length of the assessment.

Indicators of disengagement are very useful in two respects. On the one hand, they convey important information on the respondent and can be used to proxy a variety of individual traits (such as conscientiousness or the ability to endure fatigue) that are likely to influence real-life outcomes. On the other hand, these traits are not part of the skills that cognitive assessments typically try to measure. As a result, the presence of disengagement (or any kind of difference in the effort respondents exert during an assessment) biases the results of assessments and can make comparison of results across countries problematic. In this sense, information on the extent of disengagement is a useful complement to actual estimates of proficiency and can be used to make more accurate comparisons across countries.

© OECD 2019

The bottom line

Log files from computer-based assessment can help better understand respondents' behaviours and cognitive strategies. Analysis of timing information from PIAAC reveals large differences in the time participants take to answer assessment items, as well as large country differences in the share of respondents that spent the minimum amount of time necessary to understand an item. These differences in the degree of effort exerted by respondents provide a useful complement to the analysis of international skills differences.

> **CONTACT:** Marco Paccagnella (Marco.Paccagnella@oecd.org); (edu.piaac@oecd.org)

> **FOR MORE INFORMATION:**

OECD (2019), *Beyond Proficiency: Using Log Files to Understand Respondent Behaviour in the Survey of Adult Skills*, OECD Skills Studies, OECD Publishing, Paris, https://doi.org/10.1787/0b1414ed-en.

Goldhammer, F., et al. (2016), «Test-taking engagement in PIAAC», OECD Education Working Papers, No. 133, OECD Publishing, Paris, https://doi.org/10.1787/5jlzfl6fhxs2-en.

> **VISIT:** www.oecd.org/skills/piaac
Education Indicators in Focus - PISA in Focus - Teaching in Focus

The Survey of Adult Skills is a product of the OECD Programme for the International Assessment of Adult Competencies (PIAAC).

This paper is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and the arguments employed herein do not necessarily reflect the official views of OECD member countries.
This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.
You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for commercial use and translation rights should be submitted to rights@oecd.org.
The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.