OECD Education Working Papers No. 279

Developing measures of engagement in PISA

Janine Buchholz,

Marta Cignetti,

Mario Piacentini

https://dx.doi.org/10.1787/2d9a73ca-en

OECD

Organisation for Economic Co-operation and Development

EDU/WKP(2022)17

Unclassified **English text only**

21 October 2022

DIRECTORATE FOR EDUCATION AND SKILLS

DEVELOPING MEASURES OF ENGAGEMENT IN PISA

Janine BUCHHOLZ, Marta CIGNETTI, Mario PIACENTINI

OECD Education Working Paper No. 279

This working paper has been authorised by Andreas Schleicher, Director of the Directorate for Education and Skills, OECD.

This version of the document cancels and replaces the version uploaded previously. The only change implemented is the reclassification of the document, from For Official Use to Unclassified.

Janine BUCHHOLZ, Janine.BUCHHOLZ@oecd.org

JT03505685

OECD EDUCATION WORKING PAPERS SERIES

Comment on the series is welcome, and should be sent to edu.contact@oecd.org.

This working paper has been authorised by Andreas Schleicher, Director of the Directorate for Education and Skills, OECD.

---

www.oecd.org/edu/workingpapers

Abstract

Student engagement is of central importance in a low-stakes assessment such as OECD Programme for International Student Assessment (PISA). In the theoretical section, this report provides an overview of the literature on the topic and identifies sources of information to compute measures of disengagement in PISA. In the empirical section, the consistency, associations with student variables, and stability over time are examined for a set of measures based on data from PISA 2018, 2015, and 2012. The various measures investigated only show little consistency, an exception being rapid guessing on the test and non-response in the questionnaire. Boys, socio-economically disadvantaged students, as well as students with an immigrant background are more likely to show disengagement. Furthermore, disengagement is consistently associated with lower test performance. The report concludes with a discussion of possible solutions to address the impact of disengagement on the inferences made on the basis of PISA test and questionnaire data.

Acknowledgements

The analysis in this report builds on previous work conducted for the working paper by Borgonovi, Ferrara and Piacentini (2020[1]). The authors thank Alessandro Ferrara and Elodie Andrieu for support with the computation of indicators, and Tomoya Okubo for his peer review. The authors also thank their colleagues Francesco Avvisati, Ava Guez and Luis Francisco Vargas-Madriz for reviewing the manuscript during the various stages of writing this report.

PISA collects reliable and comparable data from participating countries and territories. Following OECD data regulations, a visual separation between countries and territories has been used in all charts to reduce the risk of data misinterpretation.

In 2018, some regions in Spain conducted their high-stakes exams for tenth-grade students earlier in the year than in the past, which resulted in the testing period for these exams coinciding with the end of the PISA testing window. Because of this overlap, a number of students were negatively disposed towards the PISA test and did not try their best to demonstrate their proficiency. Although the data of only a minority of students show clear signs of lack of engagement (see PISA 2018 Results Volume I, Annex A9), the comparability of PISA 2018 data for Spain with those from earlier PISA assessments cannot be fully ensured.

Data for Viet Nam are not included in tables, figures and texts that report comparisons of performance with other countries and economies' or over time, because full international comparability of results could not be assured at the time this report was published (see Annexes A4 and A6 of Volume I, PISA 2018 Results, https://doi.org/10.1787/5f07c754-en).

Table of contents

FIGURES

TABLES

ANNEX TABLES

1. Overview

In an ideal world, study participants sit an assessment by working through the material as instructed, in a conscientious and consistent manner over the whole duration of the assessment, thus showing what they know and can do and providing valuable information about their learning experiences, attitudes and beliefs. As will be outlined below, such an ideal world does not necessarily hold in practice, particularly for low-stakes assessments such as the Programme for International Student Assessment (PISA), which poses a threat to the validity of score interpretations. The examination of engagement in this context has therefore become a topic of research interest, as acknowledged by the PISA Governing Board's (PGB) decision to cover the topic in the Research, Development and Innovation (RDI) programme 2021/2022.

Although the ideal assessment situation described above is characterised by the presence of *engagement*, the phenomenon is mainly studied through the lens of *disengagement*, i.e. response behaviour that is inconsistent with behaviour expected from engaged respondents. This working paper follows this perspective, and it defines "disengaged responses" as "instances in which the individual enters a response without drawing on his or her knowledge, skills, abilities, attitudes, dispositions, and so on" (Soland, Wise and Gao, 2019, p. 151[2]). Accordingly, this broad definition covers students' disengagement when responding to both the test and the questionnaire – the two building blocks of the PISA assessment.

1.1. Disengagement in PISA tests and questionnaires: Relevance, Consequences and Correlates

1.1.1. Relevance

Standardised achievement tests and self-report questionnaires are common tools to assess what students know and can do, as well as to collect information on the context of their learning. International large-scale assessments (ILSAs) regularly administer tests and questionnaires to assess the quality of students' learning and make inferences about that of a country's educational system as a whole. In the case of PISA, country scores and rankings on the triennial assessment are commonly regarded as yardsticks to measure and compare the efficiency and effectiveness of the participating countries and economies in preparing students for their future challenges, and often drive policy change in response to the results. In some cases, the test results differed largely from the expectations of policy makers and the general public (Davoli and Entorf, 2018[3]).

An important assumption underlying the interpretation of PISA (as well as other national and international large-scale assessments) for educational policy-making lies in the claim that scores reflect differences in the targeted skills alone, and can therefore be regarded as measures of these underlying, unobservable constructs. The validity of such inferences made on the basis of the assessment, however, has to be demonstrated.

Assessment situations might not represent the ideal scenario in which all students put their best effort in their responses, as would be desired. Rather, motivation and effort have been shown to vary depending on, among other factors, the stakes attached to the testing situation. International large-scale assessments represent *low-stakes* situations to students as no personal consequences are attached. Students receive no feedback regarding their performance or any other form of incentive, and their academic record is not affected based on their performance either. It has been shown repeatedly that low-stakes assessments elicit lower levels of engagement (Wolf and Smith, 1995[4]; Finn, 2015[5]). Lower levels of engagement, in turn, are associated to lower performance in tests (Kuhfeld and Soland,

2020[6]; Pintrich and De Groot, 1990[7]; Wise and Demars, 2005[8]). It can therefore be concluded that the results of low-stakes assessments represent a mixture of cognitive and motivational factors – which ultimately causes the interpretation of results to be less straightforward. Indeed, Zamarro and colleagues (2019[9]) showed that effort explained 32 to 38% of variation in the PISA 2009 test scores. Similarly, Eklöf and colleagues (2014[10]) demonstrated that the country ranking of TIMSS (Trends in International Mathematics and Science Study) results changes when effort is statistically controlled for. Based on PISA 2015 data, Akyol and colleagues (2021[11]) estimate that a country could potentially improve its ranking by up to 15 places if all of its students had taken the test seriously. Finally, Pohl and colleagues (2021[12]) propose to disentangle different aspects of test performance such as students' propensity to omit items. The authors demonstrate how composite scores reflecting these different aspects of test behaviour would change PISA country averages, depending on the weight assigned to the different components. Summing up, the results of achievement tests in low-stakes large-scale assessments such as PISA can be regarded as a product of students' "skill and will" (Eklöf, 2010, p. 345[13]).

While PISA is considered low-stakes to students, there is much at stake for countries. As described by Akyol and colleagues (2021[11]), policy-makers in several countries anxiously await the latest ILSA results, because these provide an objective metric for the civil society to judge how well their government is doing in education. It comes as no surprise that evidence exists according to which some countries made an effort to systematically improve their test results by promoting effort on the test. For example, Canadian school teachers were given a handbook, urging them to "encourage (students) to take the assessment seriously and strive for excellence" (Akyol, Krishna and Wang, 2021, p. 186[11]). It can be argued that if all countries were to encourage their students to take the exam seriously, the change in rankings would be small. However, the effect of encouragement on test engagement is not necessarily the same across countries: Gneezy and colleagues (2019[14]) examined the change in performance when offering students with tangible monetary incentives right before starting the test. They found that while students in the US showed an increase in their performance when offered the incentive, no such effect could be observed in Shanghai.

1.1.2. Consequences

As discussed above, the most obvious consequence of disengaged responses is an underestimation of the respondent's true ability. In addition, when disengagement does not occur at random but is systematic to groups of respondents, then comparisons between these groups in terms of ability are confounded with engagement – thus posing a threat to the validity and fairness of such group comparisons.

While the discussion above was restricted to the impact of disengagement on test scores, inferences from questionnaire data are also affected in several, interrelated ways. Statistical coefficients, i.e. the inter-item correlations and internal consistency, were found to be inflated or attenuated, depending on the pattern of disengagement in the data; similarly, the first component eigenvalue is affected, meaning that disengagement holds potential to distort the factor structure (DeSimone et al., 2018[15]). Also, bivariate relationships between different scales were found to be inflated when disengagement was present in the data. In fact, it has been demonstrated that even small proportions of disengaged respondents in the dataset (5-10%) can cause spurious relationships among otherwise uncorrelated variables (Huang, Liu and Bowling, 2015[16]). For disengagement in both the test and questionnaire, another consequence affects the precision of item parameter estimates, and, thus, the estimation of scale scores in general. Finally, just as it is the case for tests, group

comparisons are confounded with disengagement when disengagement does not occur at random, but certain groups are more likely to exhibit disengagement.

1.1.3. Correlates of disengagement

To examine inter-individual differences in engagement, expectancy-value theory (EVT; e.g. Eccles and Wigfield (2002[17])) provides a useful theoretical foundation. Engagement can be conceptualised as the context-specific manifestation of motivation in terms of actions, e.g. an examinee's behaviour in a testing situation (Fredricks and McColskey, 2012[18]; Asseburg and Frey, 2013[19]) According to EVT, motivation is a function of expectancy for success and the perceived value of a task. The expectancy for success, in turn, depends on the fit between person ability and item difficulty: students of a higher ability should therefore have a higher expectancy for success while less able students should have a lower success expectancy. As a result, EVT would predict higher levels of engagement for students of higher ability, and lower engagement for less able students. Empirical findings corroborate this relationship between ability and (dis-)engagement. For example, Soland and colleagues (2019[2]) found a negative relationship between test performance and disengagement in a PISA-related assessment[1]. This is particularly noteworthy as the disengagement measure in their study was based on behaviour in the questionnaire. This finding points to the fact that disengagement is stable across, at least, the duration of the assessment situation in which both a test and a questionnaire are administered.

More evidence points to the fact that disengagement does not occur purely at random, but is associated to characteristics of the respondent. Much research has examined the relationship with gender, and found that male respondents tend to show higher levels of disengagement (e.g. Demars, Bashkov and Socha (2013[20]), Maniaci and Rogge (2014[21]), Wise, Kuhfeld and Soland (2019[22])). DeMars and colleagues (2013[20]) discuss possible explanations for this consistent finding, and point at gender differences in personality traits as well as with respect to academic motivation, all of which are likely related to disengagement. Other findings corroborate the association of disengagement and personality factors: conscientiousness, agreeableness, extraversion, and emotional stability, as reported by respondents' acquaintances, were each negatively related to disengagement (Bowling et al., 2016[23]). Bowling and colleagues (2016[23]) also found a negative relationship with class truancy, as well as rank-order consistency across time and across research situations. Taken together, it leads the authors to conclude that disengaged responding "is, in part, a reflection of enduring individual differences" (Bowling et al., 2016, p. 225[23]). Others have, however, pointed at situation-specific states such as boredom (Asseburg and Frey, 2013[19]) or affective reactions towards the test (Pintrich and De Groot, 1990[7]).

1.2. Measuring disengagement

The focus of this report and the research project is on disengagement occurring during the administration of PISA, using measures that are feasible to compute in the context of a group-administered testing situation. Before different approaches to measuring (dis-)engagement can be considered for this specific purpose, the construct itself requires a closer look. As Azevedo (2015, p. 84[24]) puts it,

> *Engagement is one of the most widely misused and overgeneralized constructs found in the educational, learning, instructional, and psychological sciences.*

[1] PISA for Schools in 85 schools in the US.

Much research found under the label of student engagement is beyond the scope of this report, as it conceptualises engagement as students' behaviour in their learning process in a broader sense, spanning the duration of a course or even school year, and including behaviour such as adhering to classroom norms, identification with the school, the completion of homework, regular class attendance or participating in extra-curricular activities (e.g. Fredricks and McColskey (2012[18])). At the same time, research relevant for the purpose of this report has been studied across various disciplines and under multiple labels. For example, survey research has a long tradition exploring *careless*, *effortless* or *insufficient effort responding* (Bowling et al., 2016[23]). In the field of cognitive assessments, the phenomenon has been studied as *test motivation* (e.g. (Baumert and Demmrich, 2001[25]), *persistence* (Hartig and Buchholz, 2012[26]), *cognitive fatigue* (Ackerman and Kanfer, 2009[27]) and *participant inattention* (Maniaci and Rogge, 2014[21]) Research incorporating the use of response time information can be found through the terms *rapid guessing* and low *response time effort* (e.g. Wise (2017[28])).The various fields have informed the planning of this study.

1.2.1. Disengagement in tests and in questionnaires

Measuring disengagement does not necessarily follow the same principles for tests and questionnaires. While an objectively "true" response (correct answer) exists for tests which can be used for validation purposes, this is not the case for questionnaires: the "true" answer of a respondent regarding their attitude, value or belief is sheer unknown. Detecting disengagement in questionnaires is therefore much more challenging and "somewhat less developed relative to those for achievement tests" (Soland, Wise and Gao, 2019, p. 151[2]). Methods for detecting disengagement need to reflect the way in which disengagement manifests itself in response behaviour, which constitutes another distinction between tests and questionnaires. Examples for disengaging in tests include gaming the system with the goal to avoid cognitive effort, for example by abusing the help functionality (Gobert, Baker and Wixon, 2015[29]). When working on questionnaires, instead, disengaged students can provide answers that, while content-relevant, do not entirely match their self-knowledge and accurately reflect their perception of themselves – this has been termed "*content-responsive faking*" (Nichols, Greene and Schmoldk, 1989, as cited in Fronczyk (2014[30])). In other instances, "*noncontent responding*" may occur, whereby respondents leave questions unanswered or respond in a random or careless way.

Different approaches to measuring students' (dis-)engagement with cognitive tests and questionnaires have been developed. Generally speaking, the approaches can be classified according to the source of information they are based on: self-report and response behaviour. As outlined in the previous paragraph, disengagement can manifest differently when responding to tests and questionnaires. Response behaviour can thus be further differentiated into test- and questionnaire-based behaviour. For each of the three categories, the next paragraphs provide an overview of methods, describe the type of information that they capture and summarise the strengths and weaknesses that accompany each of them.

1.2.2. Self-reports

Asking students directly about the level of effort they have invested when responding to the test or the seriousness of their responses in a questionnaire is not only the most obvious, but also the most common approach to measuring student engagement (Fredricks and McColskey, 2012[18]). With this approach, students are given one or several questions reflecting different aspects of engagement and are asked to select the response that best describes them. Measures based on self-reports, thus, hold the potential to reflect well-defined theoretical constructs and can be administered in a questionnaire with a relatively

short time burden for respondents (Kyllonen and Kell, 2018[31]); other advantages consist in the ease of administration, the low cost, and the feasibility of administering the question(s) in a group testing setting. However, self-reports have been shown to be sensitive to misinterpretation, lack of information, and response biases such as memory bias, social desirability bias, response style bias and reference-group bias (see Kankaraš (2017[32]) for a comprehensive review). In addition, the administration of questionnaires and self-reports, which can vary in length from just one to hundreds of items, creates somewhat of a paradox: lowly engaged students are asked to engage and spend effort to accurately respond to the presented questions – something which they are not that likely to do if they are truly disengaged (Finn, 2015[5]). Vannette and Krosnick (2014[33]) point out that accurate responses require optimisation on the respondents' part – namely, a four-step process that encompasses comprehension, retrieval, judgment and finally responding in a fashion that reflects the judgment. In absence of sufficient motivation to provide accurate answers, respondents may engage in satisficing behaviours such as straight-lining, or choosing the first available option that seems reasonable; such responses, then, are non-informative.

Measuring student engagement using self-reports has a tradition in PISA. The "*Effort Thermometer*" was first featured in PISA 2003 and administered infrequently afterwards. It is presented to students at the very end of the two-hour test and before students proceed to respond to the questionnaire. The instruction provides a reference point to be used when responding to the two questions by asking the students to think of a situation that is important to them, in which they would provide their full effort; this situation would correspond to a value of 10 on the visually represented response scale in the form of a thermometer – hence its name. Question 1, then, asks the students to indicate how much effort they put in the PISA test, using the response scale in which 10 represents full effort. Question 2 further inquiries how much effort they would have put in completing the PISA test, had they received a mark for it. In addition, the PISA questionnaire has assessed other constructs that are conceptually related and can serve as proxies of student engagement. Of particular interest for this study are the scaled indices for Work Mastery (PISA 2018) and Perseverance (PISA 2000 and 2012). Work mastery is assessed with items such as "Once I start a task, I persist until it is finished" (OECD, n.d.[34]). Similarly, perseverance is based on items such as "I remain interested in the tasks that I start" (OECD, 2014[35]). As the wording of the items for both of the indices is very general, it can be assumed to tap on behaviour that generalises across settings, including the PISA testing situation.

1.2.3. Behavioural measures

An alternative approach to measuring disengagement consists of an analysis of the behaviour exhibited by the student when working on the assessment. This can be accomplished more or less explicitly. The most explicit approach consists in the observation through a test administrator or proctor, either in a one-on-one or a group setting (an example for a standardised observation protocol for groups of students is the Baker Rodrigo Ocumpaugh Monitoring Protocol, BROMP; see Ocumpaugh et al. (2015[36])). It is intuitive to assume that disengaged response behaviour is immediately observable. Moreover, the presence of systematically trained proctors was associated to higher self-reported effort (Lau et al., 2009[37]). However, inter-rater reliability can pose a great challenge; in addition, the presence of proctors would not only be very costly and unfeasible in a group testing setting, it could potentially be perceived as intimidating and thus disrupt the response process. A less invasive approach is represented by the use of eye-tracking methodology which gathers data on fixations and saccades related to information displayed on-screen in computer-based assessments. Eye-tracking provides information on response behaviour like reading trajectories and latencies, and holds the potential to allow

for inferences about engagement. For example, eye-tracking methodology would allow to check whether a task or a questionnaire item have actually been read before an answer was given. However, the gathered data is restricted in that it only captures behaviour relative to the screen. Thus, off-screen activity cannot be classified in terms of disengagement. Off-screen activity, however, can consist in both engaged and disengaged behaviour: students might be taking task-related or task-unrelated notes using paper and pencil, they might engage in concentrated on- or off-task-related thinking while staring off-screen, or they can observe assessment-related or unrelated activities occurring in the testing room. The presence of a technical device might also be perceived as intimidating. Research on the use of eye-tracking for measuring disengagement is ongoing (Maddox et al., 2018[38]) and has not been implemented so far in the context of PISA. The remainder of this section will, therefore, focus on non-invasive measures of student engagement which are based on responses and response times that are provided through students when working through the test and questionnaire, respectively.

Questionnaire-based behavioural indicators make use of the responses and patterns thereof when responding to Likert-type matrix questions in a questionnaire. An obvious indication of disengagement is presented by a large number of missing responses. Item non-response has thus been considered to provide information on the level of effort and motivation of students (Hitt, Trivitt and Cheng, 2016[39]). Latent constructs are typically administered using matrix-type questions comprising a set of items reflecting the underlying construct. Based on the assumption that responses to items assessing the same construct should more or less be consistent, a number of indices capture the inconsistency of a response string within an individual and use it as a measures disengagement (for an overview, see Curran (2016[40])). For example, the inconsistency index (Hitt, n.d.[41]; Zamarro, Hitt and Mendez, 2019[9]) reflects the degree to which a student answers a specific item in a way that is unpredictable based on the answers to other items in the same scale. Other indices make use of semantic and psychometric pairs of antonyms/synonyms and the presence of inverse-worded items. A final set of indices for measuring disengagement indicates the degree to which students show non-differentiation (or "straight-lining") when responding to the items in a matrix question. The indices capture, more or less strictly, the extent to which a student provided the same response across a set of related questions. Such a behaviour can be the result of a careful analysis of the item content when the items are similarly hard to endorse or when a student has a very low or very high level of a particular construct. It might, however, also point at satisficing behaviour. Respondents could realise that the items in a certain set are similar and, in order to minimise the effort exerted, give the same response to all. Several studies have found support for this claim: non-differentiation is more common among less educated individuals and towards the end of a questionnaire compared to the beginning (Knowles, 1988[42]; Krosnick, 1991[43]; Vannette and Krosnick, 2014[33]), thus potentially pointing at disengagement.

Test-based behavioural measures feasible in the context of ILSAs can be based on either the responses provided to the test items or log-file information gathered in computer-based assessment (CBA). CBA is becoming more and more common among ILSAs. Tasks presented in ILSAs such as PISA typically have both open- and closed response formats and are presented on one screen each, thus rendering an analysis of response patterns as for matrix questions impossible. However, responses can either be correct or incorrect, and research has demonstrated that students differ in the degree to which they are able to maintain a consistent performance level over the duration of the assessment (e.g. Debeer et al. (2014[44]), Hartig and Buchholz (2012[26])). These differences have been shown to relate to gender (effects for socio-economic background and a discrepancy between test language and the language spoken at home were inconsistent; Wu et al. (2019[45])). However, a strong decline in performance might also be caused by increasing levels

of disengagement, for example due to fatigue or exhaustion (Borgonovi and Biecek, 2016[46]). Another approach to measuring disengagement is through behaviour captured by log-file data, i.e. information that allows to construct indices on response times and sequences of action. With the increase of computer-based testing, methods for detecting disengaged responding have also expanded (Kuhfeld and Soland, 2020[6]). A large body of research draws on response times to investigate rapid guessing (e.g. Wise (2017[28])). It is straightforward to assume that responses given below a certain time threshold cannot be considered valid, and to interpret such behaviour as the result of disengagement. Rapid guessing may be explained as a strategy adopted to complete a test within an approaching time-limit, as an indication of lack of interest, or as an indication that the student recognised he or she did not have the ability/knowledge to respond, hence they replied fast and moved on. Another promising avenue to detect disengagement consists of an analysis of the sequence of actions with which a student worked on a complex task (Gobert, Baker and Wixon, 2015[29]).

1.3. Post hoc remedies to account for disengagement

The previous sections have elaborated on the potential impact of student disengagement on the validity of inferences based on the results from tests and questionnaires, and summarised different methods for measuring disengagement in the data. Using these measures, different approaches have been proposed to account for disengagement in the data in order to mitigate the negative consequences. Generally speaking, these approaches fall into two broader categories: filtering disengaged observations from the dataset, and adjusting test scores in order to account for disengagement (Kuhfeld and Soland, 2020[6]). All of the different methods have in common that they are applied post hoc, i.e. after the data have already been collected.

Filtering consists of flagging and removing observations deemed as disengaged from the data. The removal of observations can be done either on the level of an individual response or the level of the respondent (Rios et al., 2017[47]). Removing individual responses requires an item-level indication of disengagement, something which is not possible using most of the measures introduced above as they rely on information gathered over a whole set of items. Most applications in this area make use of item-level rapid guessing (e.g. Kuhfeld and Soland (2020[6]), Rios et al. (2017[48])). Removing disengaged respondents corresponds to a listwise deletion and thus leads to a reduction of the sample size. A general issue inherent in filtering is that flagging necessarily requires the selection of one or multiple measures for detecting disengagement; however, the previous section has outlined the weaknesses that accompany each of them. Another concern relates to the fact that filtering relies on a binary decision classifying an observation as engaged or disengaged; such a decision requires setting a threshold along a continuum which marks the point at which a response or respondent can no longer be deemed engaged. Such an endeavour, just like any threshold-setting, requires a solid foundation. Finally, as outlined in the previous section, disengagement was shown to systematically relate to student characteristics such as gender. Removing disengaged observations from the dataset thus holds the potential to systematically lose observations for groups of respondents (e.g. males, low-performing students), thus distorting the representativeness of the data.

A second set of methods consists in *adjusting* the scale scores of interest (typically test scores) in order to account for disengagement in the data: one straightforward method consists in simply controlling the effect of disengagement in the scale score (e.g. Plausible Values in PISA) using regression. This approach, again, requires the selection of one or multiple measures of disengagement. Other, more elaborate methods consist of model-based adjustments in which engagement constitutes an explicit element

of the model, thus controlling its impact in the estimation of the scale score of interest. Prominent examples for such model-based approaches are Van der Linden's (2009[49]) multidimensional measurement model for both proficiency and effort, the effort-moderated Item Response Theory (IRT) model (Wise and DeMars, 2006[50]), and the speed-accuracy + engagement model (Pohl, Ulitzsch and von Davier, 2021[12]; Ulitzsch, Davier and Pohl, 2019[51]). Interestingly, these methods have in common that they all make use of response times in order to control for rapid guessing. The methodology, however, can be modified to include other measures of disengagement. Again, such an endeavour requires the selection of a disengagement measure out of the host of measures outlined above.

1.4. Aim of study

Different indicators representing three broad categories of approaches to measuring disengagement (self-report, questionnaire behaviour, and test behaviour) are feasible to compute in the context of PISA. The aim of this study is twofold. First, it focuses on the consistency between these measures as well as their relationship with student variables identified in prior research. In particular, the following will be examined based on the most recent PISA data (PISA 2018):

- The variation in disengagement between countries/economies (section 3.1);
- The consistency between the different measures of disengagement (section 3.2);
- Differences in disengagement between student sub-populations of interest – namely, gender, socio-economic status, and immigration background (section 3.3);
- The relationship between measures of disengagement and test performance (section 3.4).

Second, this study investigates how stable the various disengagement measures are across cycles, based on data from PISA 2012-2018 (section 3.5).

2. Method

2.1. Measures of disengagement

To examine disengagement in the context of PISA, three broad categories of disengagement measures have been identified in Section 1.2. Table 1 specifies which concrete measures are available or are feasible to compute for each cycle of PISA. Self-report measures were only administered in a few cycles. The questionnaire-based behavioural indicators can be computed throughout all of the cycles, regardless the mode of the assessment, as they only rely on observed response behaviour. Rapid guessing, in turn, is based on response times and can, therefore, only be computed for cycles and for countries/economies in which a computer-based assessment was administered. For PISA 2018, 69 out of 78 countries/economies administered a CBA (OECD, n.d.[34]). Each measure will be described in detail below; together they form the basis for all analyses in this report.

Table 1. Measures of disengagement by category across PISA cycles

Measure	2000	2003	2006	2009	2012	2015	2018
Self-report							
• Effort (R)		X	X		X		X
• Perseverance (R)	X				X		
• Work mastery (R)							X
Questionnaire-based behaviour							
• Non-response	X	X	X	X	X	X	X
• Non-differentiation	X	X	X	X	X	X	X
• Inconsistency	X	X	X	X	X	X	X
Test-based behaviour							
• Performance decline	X	X	X	X	X	X	X
• Rapid Guessing					X	X	X

2.2. Self-report measures

2.2.1. Effort (R)

The PISA effort thermometer, presented on the last page of the PISA assessment booklet or screen, asks students to imagine a situation which they value as important and for which they would try their best and put in as much effort as they could. It then instructs them to use this situation as a reference point, to answer two separate questions: "How much effort did you put in doing this test [PISA]?" and "How much effort would you have invested if your marks from the test were going to be counted in your school marks?" The response scale for both questions ranged from 1 to 10 with 10 being maximum effort. Only the first question is used to serve as a measure of disengagement. Answers were reverse-coded so that higher values indicate disengagement.

2.2.2. Perseverance (R)

The PISA 2012 student questionnaire contained a question (Q36) with five items asking about the extent to which student felt that the following statements described them: "When confronted with a problem, I give up easily"; "I put off difficult problems"; "I remain interested in the tasks that I start"; "I continue working on tasks until everything is perfect";

"When confronted with a problem, I do more than is expected of me". Possible answers were "Very much like me", "Mostly like me", "Somewhat like me", "Not much like me" and "Not at all like me". Based on these questions, an index based on IRT-scaling has been reported (OECD, 2014[35]). The index was reverse-coded so that higher values indicate disengagement.

2.2.3. Work mastery (R)

The PISA 2018 student questionnaire contained a question (ST182) with four items asking about the extent to which student agreed with the following statements: "I find satisfaction in working as hard as I can"; "Once I start a task, I persist until it is finished"; "Part of the enjoyment I get from doing things is when I improve on my past performance". Possible answers were "Very much like me", "Mostly like me", "Somewhat like me", "Not much like me" and "Not at all like me". An IRT-scaled index based on these questions has been reported (OECD, n.d.[34]). The index was recoded so that higher values indicate higher levels of disengagement.

2.3. Questionnaire-based behaviour

2.3.1. Non-response

The indicator captures the proportion of missing responses on all items. It is defined as:

$$Item\ non-response = \frac{Number\ of\ applicable\ items\ with\ missing\ response}{Number\ of\ applicable\ items}$$

where applicable items refers to all items included in the main background questionnaire, with the exception of questions on students' background, i.e. grade (ST001), study programme (ST002), date of birth (ST003), gender (ST004), and parental occupation (ST014, ST015). Students with non-response rates greater than 0.95 were excluded from the analysis. Higher values on the index represent higher levels of disengagement.

2.3.2. Non-differentiation

This index captures the extent to which students tend to select the same response across a set of similar and related items. It is defined as:

$$Non-differentiation = \frac{Number\ of\ applicable\ items\ sets\ without\ differentiation}{Number\ of\ applicable\ item\ sets}$$

where applicable item sets refers to questions in the PISA background questionnaire fulfilling the following conditions: (a) it is comprised of three or more items sharing the same response scale, (b) the response options represent either the degree of agreement (e.g.; "Strongly disagree") or the degree of similarity to the respondent (e.g. "Very much like me"), and (c) on the student-level, all of the items were non-missing. The number of applicable item sets could, therefore, vary between students. The final index represents the proportion of item sets in which a student did not differentiate their answers out of all the relevant items sets. Higher values on the index represent higher levels of disengagement.

2.3.3. Inconsistency

This index is based on the assumption that, in an internally consistent scale, the answer to a specific item should be correlated with the answers to the rest of the items that make up that scale. In other words, a response to a specific item in a way that is unpredictable based on the answers to other items intended to measure the same construct is considered as a

proxy of disengaged response behaviour. To construct the index, all questions in the PISA background questionnaire fulfilling the following conditions were used: (a) the question is comprised of three or more items sharing the same response scale, (b) the response options represent either the degree of agreement (e.g.; "Strongly disagree") or the degree of similarity to the respondent (e.g. "Very much like me"), and (c) on the student-level, at least three of the items were non-missing. Items were reverse-coded in case of mixed-worded scales, e.g. a negatively worded item in an item set comprised of positively worded items. Then, each item was regressed on the remaining items in its item set:

$$Y_{ijs} = \beta_0 + \beta_1 \bar{X}_{ij,-s} + \eta_{ijs}$$

where Y_{ijs} represents the answer of student i to item s in scale j, $\bar{X}_{ij,-s}$ represents the average response on all remaining items in the scale, η_{ijs} is an item-, scale- and student-specific error term representing the degree to which student i gave an unpredictable answer to item s in scale j, and β_0 and β_1 are the constant and regression weight, respectively. The index is then computed as the mean of the absolute values of η_i, thus representing average inconsistency in a student's answers across the entire PISA questionnaire. Higher values on the index thus represent higher levels of disengagement.

2.4. Test-based behaviour

2.4.1. Performance decline

This index captures the degree to which a student is able to maintain a consistent level of performance over the course of the two-hour long assessment. The idea behind this measure is that students with high levels of engagement are more likely to maintain a constant level of performance throughout the reading test while disengaged students are more likely to display a decrease in performance as a function of item position. A decrease in the performance level can thus be interpreted as the result of disengagement, but can also be related to students experiencing fatigue. The interpretation of this indicator is thus less straightforward than for the other ones. The examination of such a position effect is possible due to the test design of PISA, in which each item is presented at each cluster position, thus disentangling the confounding effect of item content and item position.

To construct the index, the following linear random coefficients model was estimated:

$$y_{ij} = \delta_0 + \delta_0^i + \delta_1 Q_{ij} + \delta_1^i Q_{ij} + \gamma_j + \theta_j + \varepsilon_{ij}$$

where y_{ij} equals zero if respondent i answered incorrectly to question j and 1 if he or she answered correctly or received partial credit. Q_{ij} represents the position of question j, rescaled for each student such that the first item in the first cluster of the 2-hour assessment takes a value of zero and the last item in the last cluster takes value 1. δ_0 represents the average student's performance on the first item in the test and δ_1 is the average performance drop from the first item to the last. γ_j is an item fixed effect to control for question difficulty and nature. θ_j are booklet fixed effects to control for the sequence of clusters in the booklet. δ_0^i is a random intercept and δ_1^i is a random coefficient that allows for students to deviate from the average values.

This procedure is applied to an item-level dataset which, for each student, includes all the items he/she attempted during the test and their sequence. The 2018 cycle of PISA was different from the previous ones because it adopted a Multi-Stage adaptive testing design. This means that the sequence of items for each student is not fully pre-determined, but depends on his/her performance at two stages of the test. In order to compute the index, the

sequence of items each student received had to be re-constructed, using information on their performance from the PISA data and information on the test design available in the technical report (OECD, n.d.[34]). Higher values on the index represent higher levels of disengagement.

2.4.2. Rapid guessing

This index is based on the assumption that answers provided below a certain time threshold are representative of rapid guessing, as opposed to response behaviour (Wise, 2017[28]; Wise, 2019[52]), and can thus be interpreted as the result of disengagement in the test. The index captures the proportion of items for which the response time is below a threshold T,

$$Rapid\ guessing = \frac{number\ of\ items\ with\ response\ times\ below\ T}{number\ of\ items}$$

where the threshold T was set to be 5 second for all items, acknowledging that alternative thresholds are equally feasible and that more fine-grained methods for setting item-specific thresholds exist (see Section 1.2.3). The threshold was set at 5 seconds with the typical PISA tasks in mind, for which it seems unrealistic to provide a meaningful answer below this threshold. Response times relate to the total time spent on an item[2].

2.5. Data

The analyses are based on the published PISA data, including students' responses and response times on the cognitive assessment as well as students' responses to the international student questionnaire, excluding national optional questionnaires. Students who took the Une Heure (UH) booklet have been excluded from the dataset as the reduced testing time and characteristics of this subsample could potentially be related to disengagement, thus confounding the effects. The resulting samples consist of $N = 612{,}004$ students in 80 countries/economies for PISA 2018, $N = 519{,}334$ students in 73 countries/economies for PISA 2015, and $N = 271{,}323$ students in 43 countries/economies for PISA 2012.

2.6. Analysis

Both the computation of the disengagement measures as well as the analyses documented in chapter 3 were conducted in STATA, version 17.0 (StataCorp, 2021[53]). Details on the specific analysis are documented in the respective section. All analyses are in line with the practice regarding the estimated standard errors accounting for the complex design of PISA. R (R Core Team, 2021[54]), version 4.1.2, has been used to produce the graphs.

[2] The PISA Public Use Files contain two sets of response time variables: time spent on the last visit, and total time across all of the visits. For details, see Annex K in the PISA 2018 Technical Report (OECD, n.d.[34]).

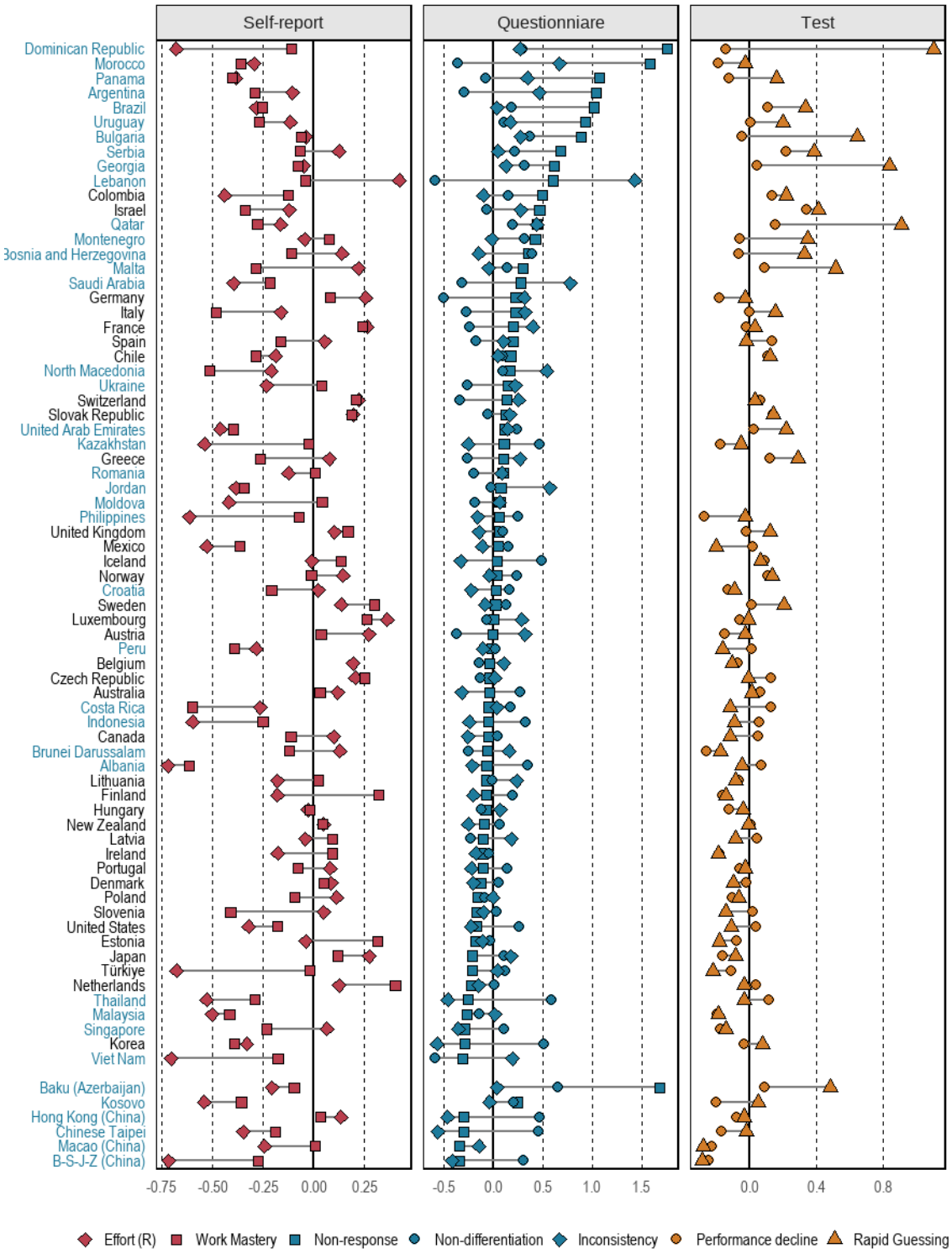
3. Results

3.1. Variation of disengagement across countries/economies

This section aims to examine the variation of disengagement across countries/economies in PISA 2018, drawing on the information provided by the various measures of disengagement. To do so, each index was aggregated on the country level. The tables in Annex A contain the comprehensive set of means and standard errors for each of the 78 countries/economies and seven disengagement measures. In addition, Figure 1 shows the distributions for all measures, separated by source of information: self-report, questionnaire-based and test-based behaviour. The countries/economies on the y-axis are ordered according to the average non-response rate, and this order was held constant throughout.

Regarding the non-response rate, the distribution of countries/economies around the OECD mean is not symmetrical: while the number of countries/economies above and below the OECD mean is about the same, countries/economies on the higher end of the distribution (indicating disengagement) show much larger deviations from the OECD mean than countries/economies on the lower end. This is particularly driven by a set of non-OECD countries/economies that exhibit non-response rates of about one and more standard deviations from the OECD mean, indicating comparatively large disengagement. The set of countries/economies forming the lower end of the distribution also consists almost exclusively of non-OECD countries/economies. Taken together, both tails of the distribution for non-response rate are comprised almost exclusively of non-OECD countries/economies. The distribution for non-response rate shows another interesting pattern: the nine countries/economies on the low end of the distribution, exhibiting the lowest non-response rates overall, are exclusively located in East and Southeast Asia. Non-response rate as a measure of disengagement would thus lead to conclude that students in these countries/economies are, on average, more engaged when working on the questionnaire. This finding does not hold when inspecting differences in self-reported engagement as measured through the effort thermometer: in some Asian countries/economies (e.g. Korea) students report a higher level of engagement than the OECD average, while in some other (e.g. Japan) they report a lower level of engagement. Latin American countries, such as Colombia and Dominican Republic, exhibit relatively large levels of disengagement according to behavioural measures such as non-response and rapid guessing, yet their students report relatively high levels of engagement with the PISA test. This finding reflects a more general pattern: in Figure 1, the distribution of country means is unsystematic, meaning that the order of countries/economies according to average non-response is rather independent from the order on the other two measures. This pattern holds across all of the seven measures. None of the countries/economies exhibits a low or high level of disengagement throughout the whole set of measures. This finding points at a low consistency among the measures, which is subject to the next section (Section 3.2).

Figure 1. Distribution of disengagement for selected measures



Notes: The analyses are based on data from PISA 2018. All indices are coded so that higher values indicate higher levels of disengagement; the indices were standardised to have a mean of 0 and standard deviation of 1 across the OECD student population. Rapid guessing is only available for countries/economies that administered computer-based assessment. B-S-J-Z (China) refers to the four PISA-participating provinces/municipalities of the People's Republic of China: Beijing, Shanghai, Jiangsu and Zhejiang.
Countries and economies are ranked in descending order of the level of non-response.

3.2. Consistency among measures of disengagement

This section examines the consistency among the different measures of disengagement, both within and between the three broad categories (self-report, questionnaire-based and test-based behaviour) based on the PISA 2018 data. The analyses have been conducted both on the individual and on the country level.

On the individual level, the Pearson correlation for each pair of the seven measures has been computed. The results are shown in Table 1. The most striking finding relates to the fact that most of the coefficients are small in absolute size, and some are even negative. This finding points at a low consistency among the measures: whether a student is deemed as disengaged on one measure is only loosely associated to the finding on another measure. Even worse, in case of negative relationships, students deemed disengaged on one measure are considered engaged on the other. Table 1 has been arranged according to the three broad categories of measures.

- The two *self-report* measures, effort and work mastery, exhibit a small positive correlation, showing at least some degree of consistency. Students who report lower levels of effort tend to report lower levels of work mastery, too.
- Within the three *questionnaire-based* behavioural measures, however, the findings are fairly mixed. While non-response is lowly correlated to both non-differentiation and inconsistency, a strong negative correlation exists between the latter two. Students deemed as disengaged based on non-differentiation would be deemed rather engaged based on inconsistency. What seems like a contradiction at first glance can be explained as follows: the negative association means that students who engage in non-differentiation behaviour are less likely to also exhibit inconsistency. Although both are based on response patterns, the two indices capture mutually exclusive patterns. For example, straight-lining response behaviour across a large set of items would lead to a high score on non-differentiation, but a low score for inconsistency. The negative relationship, thus, indicates that students engage in either one of the two behaviours, and do not alter between them.
- Finally, the two *test-based* behavioural measures, performance decline and rapid guessing, show a small positive relationship, meaning that students who engage in rapid guessing are also more likely to exhibit a decline in performance. The finding might be explained by the fact that performance decline can only occur for students with an initial performance level that was not low. Students who engage in rapid guessing behaviour may, therefore, show lower performance levels. Whether this potential explanation holds is subject to analysis in the remainder of this report (see section 3.4).

Across the set of measures, rapid guessing seems to be related to most of the others. A particularly high correlation exists between rapid guessing and non-response, meaning that students who tend to respond to test items below a reasonable amount of time are more likely to leave items in the questionnaire unanswered. The finding is interesting in that it relates behaviour on the test to that in the questionnaire.

Table 2. Consistency among measures of disengagement (individual level)

	Self-report		Questionnaire-based behaviour			Test-based behaviour	
	Effort (R)	Work mastery (R)	Non-response	Non-differen-tiation	Incon-sistency	Perfor-mance decline	Rapid guessing
Effort (R)	1.00						
Work mastery (R)	.16***	1.00					
Non-response	.06***	.06***	1.00				
Non-differentiation	-.03***	-.06***	.05***	1.00			
Inconsistency	.09***	.10***	.15***	-.49***	1.00		
Performance decline	.03***	.01***	.02***	.03***	.01***	1.00	
Rapid guessing	.12***	.07***	.32***	.09***	.13***	.05***	1.00

Notes: The analyses are based on data from PISA 2018. All indices are coded so that higher values indicate higher levels of disengagement. *** $p < .001$.

The second set of analyses is conducted on the country level and thus makes use of the country averages (see Tables in Annex A). Table 2 contains the respective results: the upper triangular contains the Pearson correlation coefficients while the lower triangular contains scatterplots to allow for a visual inspection of the respective relationship, separately for OECD- and non-OECD countries/economies.

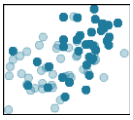
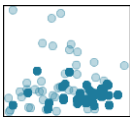
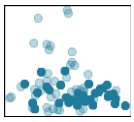
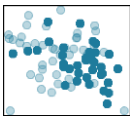
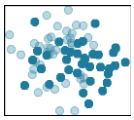
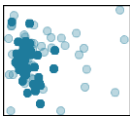
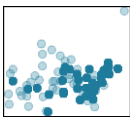
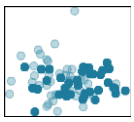
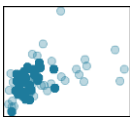
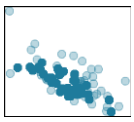
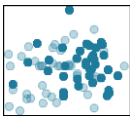
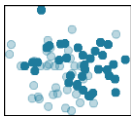
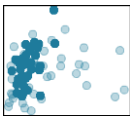
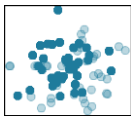
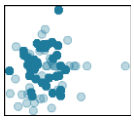
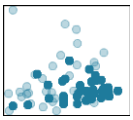
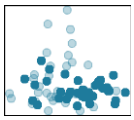
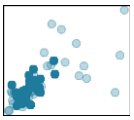
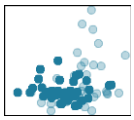
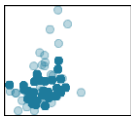
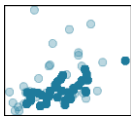
Just as on the individual level, some of the correlation coefficients are negative. Such negative relationships mean that there are countries/economies in which students tend to show disengagement on measure A, but not on measure B, while the in other countries/economies, students are likely show disengagement on measure B but not on measure A. Within each group of measures, the direction for the relationship between the measures is the same as on the individual level, albeit more pronounced. Also, the strong negative relationship between non-differentiation and inconsistency persists, meaning that students in any given country share their tendency to engage in either one or the other response pattern. In addition, rapid guessing shows strong correlations with all other measures, except for self-report measures: in a given country, the tendency of students to engage in rapid guessing on the test is only loosely related to whether students also report disengagement. This finding might point to general, culture-related response biases; that is, response tendencies that are unrelated to item content (e.g. van de Vijver and He (2014[55])). Finally, a strong positive relationship between rapid guessing and non-response is equally found on the country-level, meaning that countries/economies in which students engage in rapid responding are more likely that their students leave questionnaire items unanswered.

The low correlations between measures could suggest that the indicators measure different aspects of engagement and/or suffer from different sources of bias or measurement error.

Table 3. Consistency among measures of disengagement (country level)

Pearson correlation coefficients (upper triangular), scatter plots (lower triangular)

	Self-report		Questionnaire-based behaviour			Test-based behaviour	
	Effort (R)	Work mastery (R)	Non-response	Non-differentiation	Inconsistency	Performance decline	Rapid guessing

Effort (R)	1.00	.49†	-.05	-.29	.19	-.02	.04
Work mastery (R)		1.00	-.17†	-.11	-.08	-.08	-.10†
Non-response			1.00	.02	.48	.16	.70*
Non-differen-tiation				1.00	-.66*	.08	.26
Inconsistency					1.00	.11	.38
Performance decline						1.00	.38*
Rapid guessing							1.00

Notes: The analyses are based on data from PISA 2018. All indices are coded so that higher values indicate higher levels of disengagement. *$p < .05$, †$p < .10$. Points in darker colour represent OECD countries.

3.3. Differences in disengagement between groups

This section examines whether differences in disengagement are systematically related to student characteristics, drawing on the data from PISA 2018. In particular, this section will analyse the association of disengagement with gender, socio-economic status, and immigrant background. The analyses are conducted on the individual level, and group differences on the standardised measures will form the core of the results.

Figure 2 shows the respective findings for *gender*, i.e. the difference between girls and boys on the seven disengagement measures. In addition, the tables in Annex B contain the comprehensive set of results, i.e. the difference scores as well as associated standard errors. Negative differences indicate higher disengagement for boys. Indeed, the differences are almost exclusively negative across all countries/economies and all measures. The few positive results (indicating higher disengagement for girls) are almost exclusively non-significant. The only exceptions for significantly higher disengagement for girls are the following:

- Work mastery (R) for Korea
- Non-response in Lebanon
- Non-differentiation in Mexico, Lebanon, the Philippines, and Romania

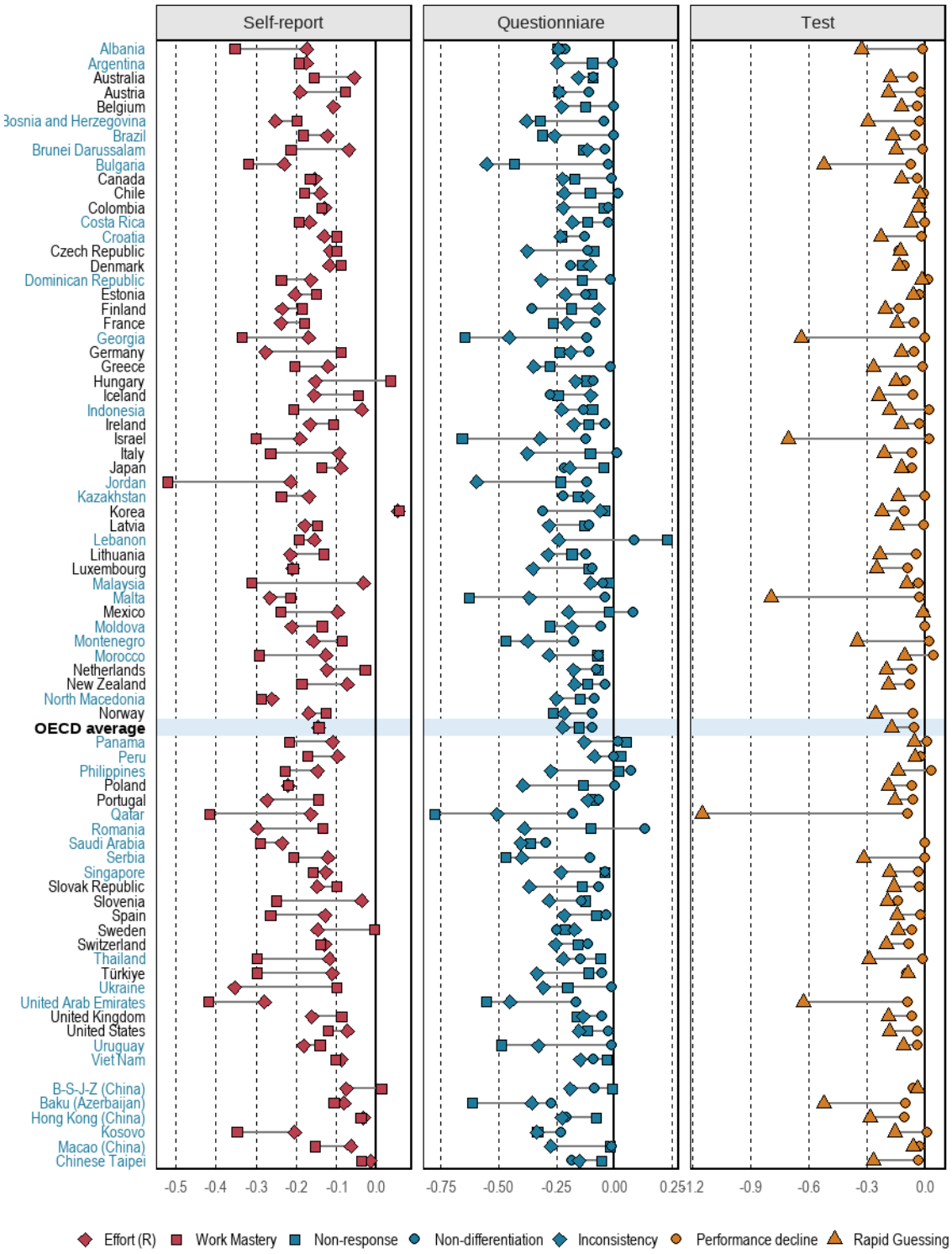
- Performance decline in Morocco and the Philippines

In all of these instances, the differences are small in magnitude. This finding, thus, provides overwhelming evidence in support of previous research findings, according to which males more likely to exhibit disengagement.

For *socio-economic status*, the analyses are based on the two groups constituting the top and bottom quartile on the index of social and economic status (ESCS) within each country, respectively. Figure 3 shows the respective findings; the tables in Annex C contain the comprehensive set of results, i.e. the difference between the two groups as well as associated standard errors. Negative differences (top quarter - bottom quarter) indicate higher disengagement for the most disadvantaged students (bottom quarter), while positive values mean lower disengagement for the most advantaged students. The results show negative coefficients for the vast majority of countries/economies on most of the measures, particularly work mastery (R), non-response, inconsistency, and rapid guessing. Disadvantaged students are thus clearly less engaged than advantaged students, suggesting that differences in the capacity to engage and sustain effort magnify the socio-economic gap in performance. Part of the observed gap in the rapid guessing can be explained by the fact that disadvantaged students, who tend to perform worse on the test, are more likely to disengage because they find the test too difficult. Findings for effort, non-differentiation and performance decline are somewhat mixed.

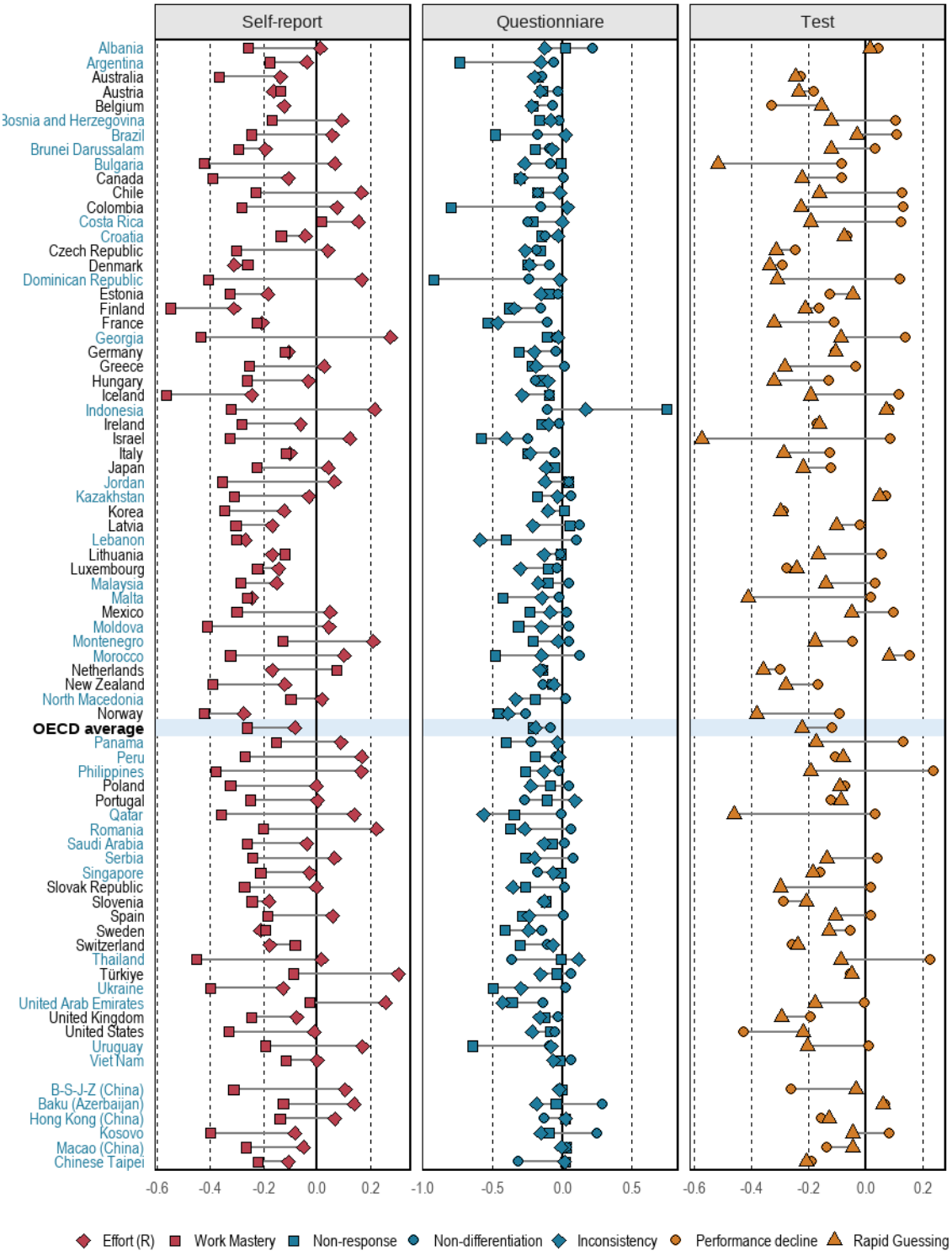
For *immigrant background*, the analyses are based on the published index IMMIG (Index Immigration status). It has three levels to differentiate native, second generation, and first-generation students. The latter two categories have been collapsed, resulting in two groups of students (with and without an immigrant background). Figure 4 shows the respective findings; in addition, the tables in Annex D contain the comprehensive set of results, i.e. the difference between students with and without an immigrant background, as well as the associated standard errors. Positive differences indicate higher disengagement for students with an immigrant background. For the vast majority of countries/economies and on most of the measures, the differences are positive, indicating higher levels of disengagement for students with an immigrant background. The results are less clear and somewhat mixed for non-differentiation and performance decline.

Figure 2. Differences in disengagement by gender (girls – boys)



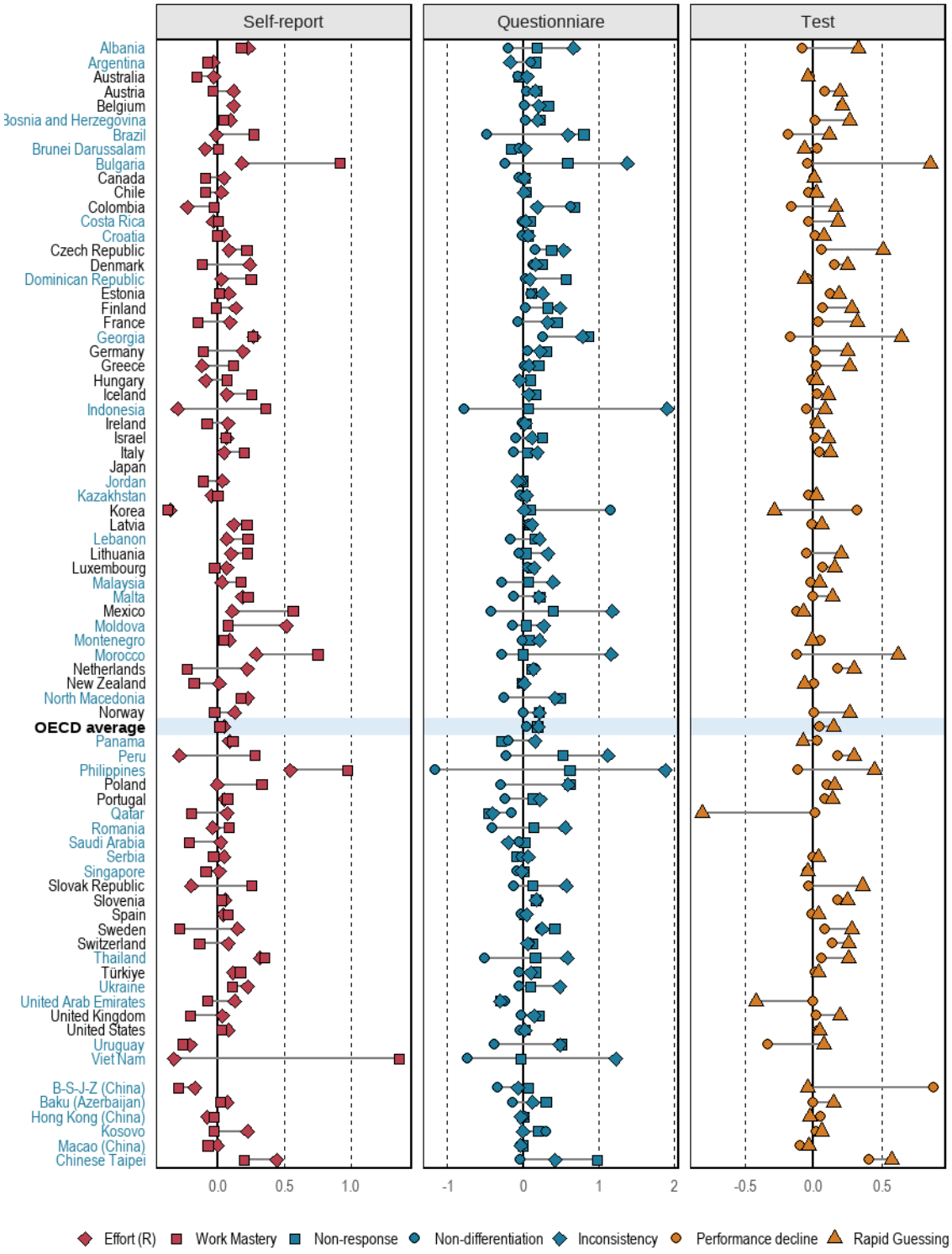
Notes: Values represent differences between OECD standardised measures (mean and standard deviation of 0 and 1, respectively, across OECD countries). Negative differences (girls-boys) indicate higher disengagement for boys. Rapid guessing is only available for countries/economies that administered computer-based assessment.

Figure 3. Differences in disengagement by socio-economic status (top quarter - bottom quarter)



Notes: Values represent differences between OECD standardised measures (mean and standard deviation of 0 and 1, respectively, across OECD countries). Negative differences (top quarter - bottom quarter) indicate higher disengagement for the bottom quarter on the ESCS. Rapid guessing is only available for countries/economies that administered computer-based assessment.

Figure 4. Differences in disengagement by immigration background (non-immigrant - immigrant background)



Notes: Values represent differences between OECD standardised measures (mean and standard deviation of 0 and 1, respectively, across OECD countries). Negative differences (non-immigrant - immigrant background) indicate higher disengagement for students with an immigrant background. Rapid guessing is only available for countries/economies that administered computer-based assessment.

3.4. Relationship of disengagement with performance

This section examines the relationship between disengagement and test performance. As Reading presented the major domain in PISA 2018, test performance is represented by the Plausible Values for Reading only. To address the research question, regression analyses of the PISA Reading Plausible Values on each of the engagement measures[3] while controlling for student background characteristics have been conducted. Control variables were gender (ST004D01T), socio-economic status (ESCS), immigrant background (IMMIG, for which second- and first-generation have been collapsed), age, grade (i.e. grade compared to modal grade in country), and whether the language spoken at home most of the time (variable ST022Q01TA) was different from the test language (LANGTEST_COG). The regression weight for a measure of disengagement thus represents the change in reading performance (on average across the Plausible Values) on the PISA scale that is associated to an increase of one standard deviation on the respective disengagement measure (all measures of disengagement were standardised such that their mean and standard deviation are 0 and 1 across the OECD student population, respectively).

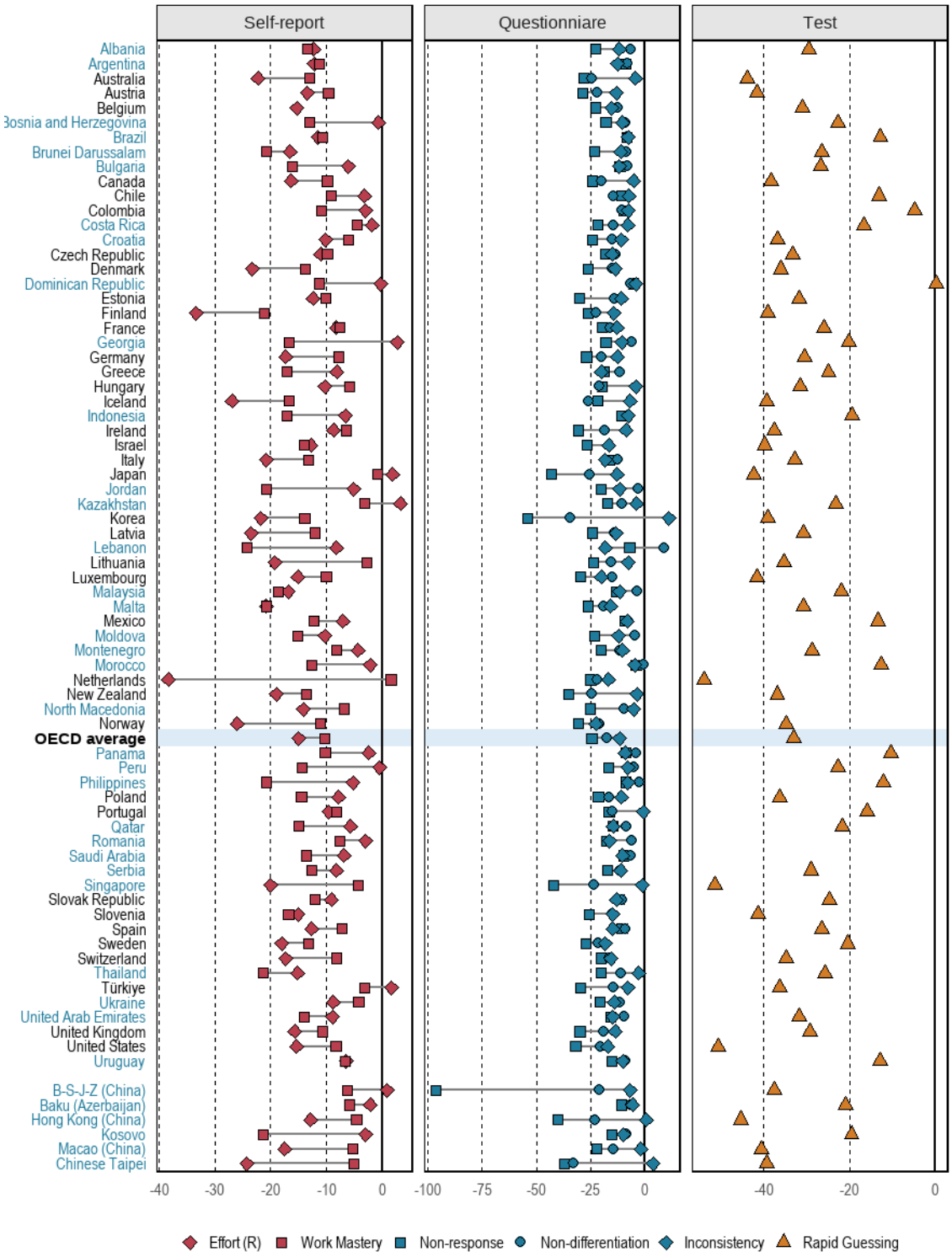
For each of the seven disengagement measures, the tables in Annex E contain the full set of regression estimates (weights) and associated standard errors for all countries/economies. The vast majority of weights are negative and significantly different from zero. In line with expectations, the findings show that each measure of disengagement is associated with lower performance on the test. The few exceptions, i.e. significant positive weights, occurred for:

- Self-reported effort (R) in Georgia and Kazakhstan
- Non-differentiation in Lebanon
- Inconsistency in Korea and Chinese Taipei

Among the set of measures, rapid guessing and non-response show the largest effects, meaning that deviations from the OECD mean on these two measures are associated to larger differences in performance on the test, compared with the other measures. Figure 2 further illustrates the findings for a selected set of measures, i.e. self-reported effort, non-response and rapid guessing. Overall, the different disengagement measures seem to pick up on behaviours and traits that are related to test performance. The finding is also interesting in that it connects behaviour on the test (performance) to measures based on information collected after the administration of the test, i.e. self-reports and response patterns collected with the questionnaire. The magnitude of the relationship is quite high for most countries/economies, although significant differences are observed across countries/economies. Keeping in mind the large differences in the measured level of engagement we observe across countries/economies, the results suggest that country rankings in PISA would be affected by using cognitive measures that are adjusted to account for engagement.

[3] Performance decline was excluded from the analysis, as it draws on the same information as the performance measure.

Figure 5. Relationship between disengagement measures and reading performance



Notes: The analyses are based on data from PISA 2018. Values represent the predicted change on the PISA Reading scale associated with a one standard deviation increase on the respective disengagement measure after accounting for student characteristics (e.g. gender and socio-economic status). Rapid guessing is only available for countries/economies that administered computer-based assessment.

3.5. Stability of disengagement across cycles

This section investigates the consistency of disengagement across cycles, based on data from PISA 2018, 2015, and 2012. As the PISA study design is cross-sectional, the analyses are based on country-level means. In particular, for each country, the raw difference between cycles for a given measure is computed, along with the standard error of this difference. Differences between cycles were based on unstandardised measures. Small differences indicate high consistency on the particular measure in the specific country. High consistency in a given country thus indicates that the particular measure picks up on something (self-report, behaviour) that is stable over the student cohorts in this country.

The following indicators form the basis of the analyses in this section:

- Self-reported effort (R), administered in PISA 2018 and 2012;
- Questionnaire non-response, based on all applicable items in 2018[4], PISA 2015[5] and 2012[6];
- Fast guessing, separately for Reading, Science and Mathematics, based on the respective subset of items administered in both PISA 2018 and 2015.

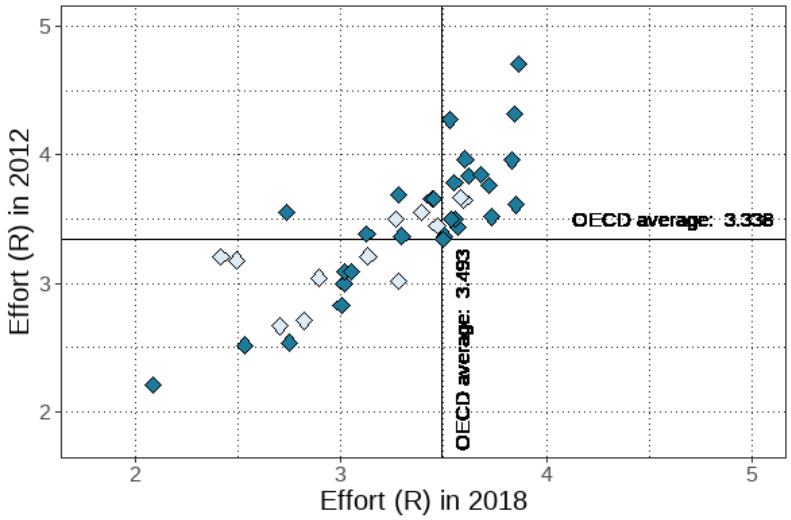
For *self-reported effort (R)*, the findings are displayed in Figure 6 and detailed information is documented in Annex F. The analyses are based on 41 countries/economies with available data in both 2018 and 2012. The within-country difference is significant in about two-thirds of the countries/economies, i.e., the mean level of self-reported effort in these countries/economies has changed over time. However, across countries/economies, the correlation of country means is $r = .81$ ($p < .001$), thus indicating a strong relationship: the tendency of students to answer to the effort thermometer in a certain way is rather consistent over the duration of six years.

[4] Applicable items in PISA 2018 include all items included in the main background questionnaire, with the exception of questions on students' background, i.e. grade (ST001), date of birth (ST003), and gender (ST004).

[5] Applicable items in PISA 2015 include all items included in the main background questionnaire, with the exception of questions on students' background, i.e. grade (ST001), date of birth (ST003), gender (ST004), education of mother (ST005 and ST006) and father (ST007 and ST008), home possessions (ST012), country of birth (ST019), life satisfaction (ST016), bullying (ST038), as well as ST059, ST060, ST061, ST065, ST071, and ST129 due to their response format.

[6] Applicable items in PISA 2012 include all items included in the main background questionnaire, with the exception of questions on students' background, i.e. grade (ST01), date of birth (ST03), age at ISCED 1 (ST06) as well as questions ST06, ST21, ST25, ST57, ST69, ST70, ST71, ST72, ST76 due to their response format.

Figure 6. Association of self-reported effort in PISA 2018 and 2012

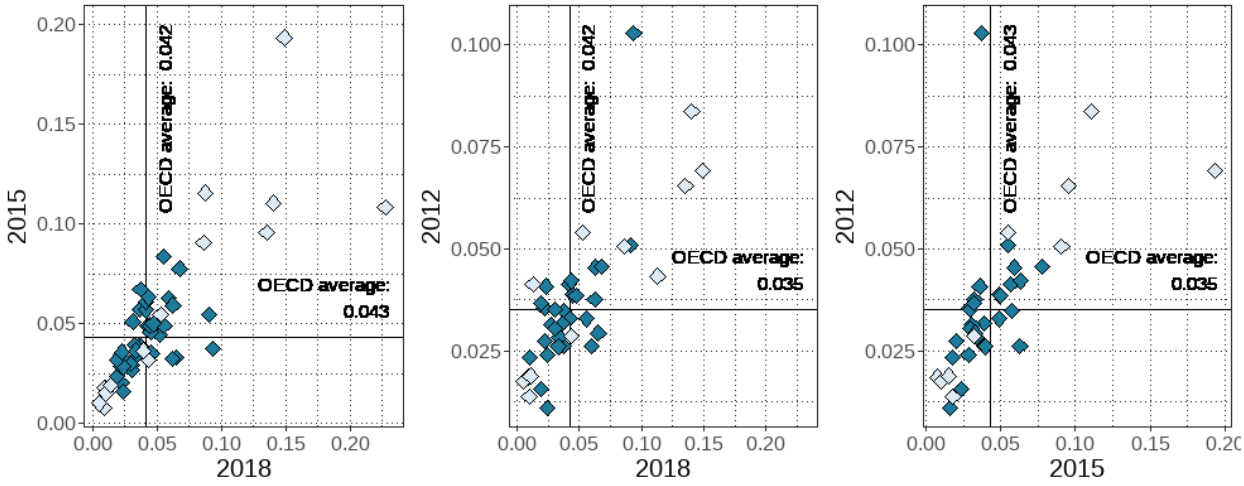


Note: Points represent country means; points in darker colour represent OECD countries.

For *non-response*, the findings are displayed in Figure 7 and detailed information is documented in Annex G. The comparisons are based on 53 (2018-2015), 42 (2018-2012), and 40 (2015-2012) countries/economies, respectively, which participated in the two respective PISA cycles.

For all three sets of comparisons, most within-country differences are significant, thus indicating a change in the absolute level of non-response rates in these countries/economies. However, strong associations exist across countries/economies: the correlations are $r = .80$ ($p < .001$, 2018-2015), $r = .78$ ($p < .001$, 2018-2012), and $r = .65$ ($p < .001$, 2015-2012). These findings demonstrate a high level of consistency over time regarding the relative tendency of students in these countries/economies to leave questions unanswered.

Figure 7. Association of non-response in PISA 2018, 2015 and 2012

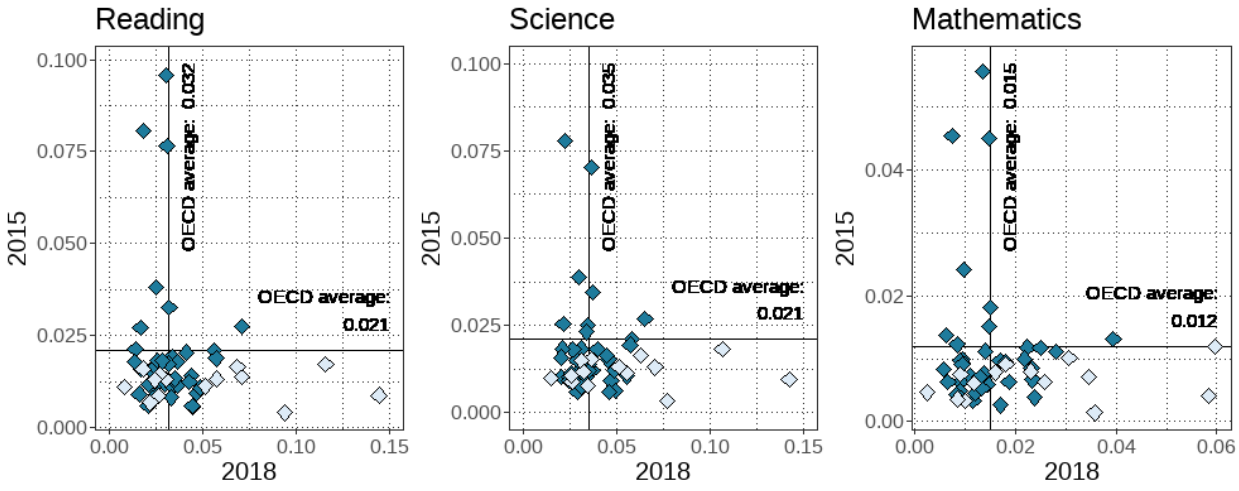


Notes: Points represent country means; points in darker colour represent OECD countries.

For *rapid guessing*, the findings are displayed in Figure 8 and detailed information is documented in Annex H. For all three domains, the comparisons are based on 52 countries/economies which administered a computer-based assessment in both 2018 and 2015.

For all three sets of comparisons, most within-country differences are significant, thus indicating a change in the prevalence of rapid guessing over time. Across countries/economies, however, the correlations are not significant: $r = -.14$ ($p = .32$, Reading), $r = -.12$ ($p = .41$, Science), and $r = -.09$ ($p = .50$, Mathematics), respectively. These non-significant correlations indicate low levels of stability: in a given country, students may have engaged in rapid responding relatively frequently in one cycle, but less so in the other. This finding is rather surprising, particularly given the high levels of stability for self-reported effort and non-response. The change in the major domain of assessment between PISA 2015 (Science) and PISA 2018 (Reading) does not provide an explanation for the instability, as the analyses were restricted to the joint set of items administered in both cycles. However, the findings need to be interpreted with caution as the distributions are extremely skewed. While most countries/economies show very low levels of rapid responding in both cycles, there are a few outliers, i.e., countries/economies in which students showed an extreme value for rapid guessing in one cycle, and not in the other.

Figure 8. Association of rapid guessing in PISA 2018 and 2015 for Reading, Science, and Mathematics



Note: Points represent country means; points in darker colour represent OECD countries.

4. Discussion

Student engagement is of central importance in the context of low-stakes large-scale assessments such as PISA. The PISA Governing Board has acknowledged this fact by implementing a project on the topic as part of the PISA Research, Development and Innovation programme. The first steps along this research line consist (1) in a review of feasible methods for measuring disengagement, (2) to implement these using PISA data, and (3) to explore the variation, correlates and consistency of those measures accordingly. This report serves to document these.

A central result regarding of the first aim the first step consists is the identification of three broad categories of disengagement measures, each representing a different source of information which can be used to inform the construction of potential measures: self-report, behaviour on the test, and behaviour on the questionnaire. This categorisation is based on an extensive literature search for studies on the phenomenon of disengagement, conducted across disciplines and using different keywords. An alternative approach would have consisted in defining the psychological state of disengagement, and finding ways to operationalise it. Such an endeavour would certainly fill a research gap, but would also extend the scope of the current research agenda. Fine-grained methods such as cognitive laboratories, observations, and a triangulation of the various methods would be required. The approach taken in this project, in contrast, is rather pragmatic by taking stock and classifying methods that have already been proposed, and by evaluating their feasibility in the context of PISA.

In order to tackle the second step, previous cycles of PISA were examined to document which of the measures are (potentially) available. As a result, seven measures of disengagement were computed for PISA 2018, and a subset of those for PISA 2015 and 2012, respectively. These measures cover the three broad categories identified earlier and are fairly diverse with respect to the information they capture. Using these measures, the third step was to examine the variability of disengagement across countries/economies, their consistency, correlates with student characteristics and performance, as well as their stability across cycles. The key results can be summarised as follows:

- The consistency between most of the measures was low. Correlations between measures, both on the individual and on the country level, were small in magnitude or even negative. The measures, therefore, seem to capture different, sometimes mutually exclusive, manifestations of behaviour or traits which can be considered to represent disengagement. The low correlation between self-report and behavioural measures suggest that some students might not be able to correctly evaluate their actual engagement, or that they might be inclined to declare they have been effortful because of social desirability. Self-report measures might be systematically influenced by response biases, particularly acquiescence. Finally, it is also possible that these measures are differently exposed to measurement error. Interestingly, rapid guessing on the test and high non-response rates in the questionnaire show a relatively strong correlation. These two measures are the most straightforward to compute and to interpret: while non-differentiated responses to the questionnaires might reflect engaged behaviour (some students might indeed ‘strongly agree’ with all the statements in a questionnaires), it is quite unlikely that an engaged student would leave most questions blank in a questionnaire, or would not dedicate enough time to read a prompt and try to answer items in the test. While more research and development of better measures is warranted, the results of the present study suggest that it would be appropriate to focus on these two measures – rapid guessing in the test and missing responses in the questionnaire – in future

analyses. The method for constructing the rapid guessing indicator can be further refined through the implementation of more appropriate (e.g. item-specific) time thresholds.

- Disengagement, measured in various different ways, is systematically related to student characteristics. In particular, boys, socio-economically disadvantaged students, as well as students with an immigrant background were more likely to show disengagement in the majority of countries/economies. The results are particularly clear with respect to gender, a finding which is in line with previous research.
- Disengagement was shown to systematically relate to lower test performance. This finding is interesting as test performance was not only related to behaviour on the test, but also to behaviour that students demonstrated in the questionnaire administered afterwards.
- Across cycles, both self-reported effort and questionnaire non-response are strongly associated, pointing at rather stable tendencies of students in these countries/economies to engage in the respective form of disengagement. In the case of self-reported effort, the strong association might be caused by stable response tendencies on the country-level, particularly acquiescence (Buchholz, 2022[56]; Johnson et al., 2005[57]). Surprisingly, rapid guessing on the test proved to be rather unstable across cycles.

The potential consequences of disengagement on PISA results warrant additional research. From a theoretical perspective, the most serious consequence consists in a systematic underestimation of students' true ability. The present analyses provided evidence for such systematic relationships between disengagement and performance. However, no causal interpretation is possible due to the cross-sectional nature of the data: it is not clear whether disengaged students chose to not put their best effort in the test, or whether low-performing students tend to exhibit behaviour that is captured by the disengagement measures. Either longitudinal designs or experimental studies, manipulating features of the assessment believed to affect engagement, would be required to establish the true magnitude of the effect of disengaged behaviour on PISA performance. Other consequences of disengagement in the data relate to technical aspects and should also be examined in more detail. In particular, future studies should examine the influence of the degree of disengagement in the data on country-level statistics and score comparability among countries/economies. This is particularly relevant as disengagement holds the potential to cause residual correlations of the items, thus violating the local independence assumption, and potentially introducing bias to country score rankings.

The next step in the research agenda consists in identifying solutions to remedy or control the impact of disengagement on the inferences made on the basis of test and questionnaire data. Several avenues for post hoc approaches have been introduced in the theoretical part of this report. All of these methods rely on the selection of one or multiple measures in order to control for the effect of disengagement. Given the low consistency between measures, the selection will immediately affect the result of the post-hoc method. The selection, therefore, cannot be arbitrary. Clear guidance should exist as to which measure is superior or should be used for a particular purpose. In addition, if such methods are applied in the context of a large-scale assessment, it is essential to keep the procedure identical across countries/economies and across cycles, as such a correction could otherwise bias the respective comparisons.

Summing up, post-hoc methods for correction do not present a promising avenue for tackling disengagement. No correction would be fully reliable, and there is a risk that such

adjustments would add more noise to the reported scores and confuse the users of the reports and of the data. Filtering does not present a good solution either, because it leads to the systematic loss of observations and, thus, to a loss of representativeness. Instead, assessments should be developed from the beginning in a way that they elicit engagement, or prevent students from becoming disengaged as they continue to work on them. Several features of assessments can be considered for this purpose, in particular:

1. Tasks that less strongly resemble classical test items, but which instead allow for exploration, interactivity, use of multimedia learning materials, and/or present game-based environments which students might associate to leisure-time experiences.
2. Particular item characteristics (e.g. response format, the amount of text in the stimuli, or the presence of a picture; both in isolation as well as in combination with the item position).
3. Allocation of test items of appropriate difficulty, as accomplished through the use of adaptive testing designs.
4. Reduction of testing time, for example through the use of adaptive testing which allows for greater measurement accuracy using fewer test items, through a more efficient design of questionnaires using routing and matrix sampling, or through the collection of data in multiple sessions or with longer test breaks.
5. The introduction of a warning message or other type of signal during an ongoing assessment whenever disengagement is detected (through real-time use of process-data).

The first three features are subject to ongoing research within the PISA Research, Development and Innovation programme. Regarding (1), innovative item designs are under development and testing for both the PISA 2025 Innovative Domain (Learning in the Digital World) and the Platform for Innovative Learning Assessments (PILA). Enhanced cognitive laboratories are being implemented for these new assessments that will collect extensive concurrent and retrospective information on engagement and related factors, such as confusion, frustration or perceptions of difficulty. An eye-tracking study is also being organised to get better insights on how students engage with these more complex and interactive test environments. For (2), an ongoing study systematically examining the effect of item characteristics and features of the assessment design using PISA data will be published in a forthcoming OECD working paper. The results of this analysis will provide guidance the development of future instruments. For (3), an ongoing study is examining the effect of booklet misallocations on engagement using PISA 2018 data. A specific feature of the multi-stage adaptive testing design in PISA 2018 allows to systematically examine the effect of a (mis-)match between item difficulty and student performance, thus providing evidence on the benefit of adaptive designs in terms of (dis-)engagement.

In parallel, the RDI project will continue to explore opportunities to derive more robust measures of engagement with PISA data. This is not much for the sake of correcting country rankings but because engagement is an important construct to measure in PISA and, arguably, a core education goal in itself.

References

Ackerman, P. and R. Kanfer (2009), "Test Length and Cognitive Fatigue: An Empirical Examination of Effects on Performance and Test-Taker Reactions", *Journal of Experimental Psychology: Applied*, Vol. 15/2, pp. 163-181, https://doi.org/10.1037/a0015719. [27]

Akyol, P., K. Krishna and J. Wang (2021), "Taking PISA Seriously: How Accurate are Low-Stakes Exams?", *Journal of Labor Research*, Vol. 42/2, pp. 184-243, https://doi.org/10.1007/S12122-021-09317-8. [11]

Asseburg, R. and A. Frey (2013), "Too hard, too easy, or just right? The relationship between effort or boredom and ability-difficulty fit", *Psychological Test and Assessment Modeling*, Vol. 55/1, pp. 92-104. [19]

Azevedo, R. (2015), "Defining and Measuring Engagement and Learning in Science: Conceptual, Theoretical, Methodological, and Analytical Issues", *Educational Psychologist*, Vol. 50/1, pp. 84-94, https://doi.org/10.1080/00461520.2015.1004069. [24]

Baumert, J. and A. Demmrich (2001), "Test motivation in the assessment of student skills: The effects of incentives on motivation and performance", *European Journal of Psychology of Education 2001 16:3*, Vol. 16/3, pp. 441-462, https://doi.org/10.1007/BF03173192. [25]

Borgonovi, F. and P. Biecek (2016), "An international comparison of students' ability to endure fatigue and maintain motivation during a low-stakes test", *Learning and Individual Differences*, Vol. 49, pp. 128-137, https://doi.org/10.1016/j.lindif.2016.06.001. [46]

Borgonovi, F., A. Ferrara and M. Piacentini (2020), "From asking to observing. Behavioural measures of socio-emotional and motivational skills in large-scale assessments", *DoQSS Working Papers*, https://ideas.repec.org/p/qss/dqsswp/2019.html (accessed on 13 October 2022). [1]

Bowling, N. et al. (2016), "who cares and who is careless? Insufficient effort responding as a reflection of respondent personality", *Journal of Personality and Social Psychology*, Vol. 111/2, pp. 218-229, https://doi.org/10.1037/PSPP0000085. [23]

Buchholz, J. (2022), "Mixed-worded scales and acquiescence in educational large-scale assessments", *OECD Education Working Papers*, No. 269, OECD Publishing, Paris, https://doi.org/10.1787/8dd310c0-en. [56]

Curran, P. (2016), "Methods for the detection of carelessly invalid responses in survey data", *Journal of Experimental Social Psychology*, Vol. 66, pp. 4-19, https://doi.org/10.1016/j.jesp.2015.07.006. [40]

Davoli, M. and H. Entorf (2018), "The PISA Shock, Socioeconomic Inequality, and School Reforms in Germany", *IZA Policy Paper*, No. 140, https://docs.iza.org/pp140.pdf (accessed on 24 March 2022). [3]

Debeer, D. et al. (2014), "Student, School, and Country Differences in Sustained Test-Taking Effort in the 2009 PISA Reading Assessment", *Journal of Educational and Behavioral Statistics*, Vol. 39/6, pp. 502-523, https://doi.org/10.3102/1076998614558485. [44]

Demars, C., B. Bashkov and A. Socha (2013), "The Role of Gender in Test-Taking Motivation under Low-Stakes Conditions", *Research and Paractice in Assessment*, Vol. 8, pp. 69-82, https://eric.ed.gov/?id=EJ1062839. [20]

DeSimone, J. et al. (2018), *The Differential Impacts of Two Forms of Insufficient Effort Responding*, Blackwell Publishing Ltd, https://doi.org/10.1111/apps.12117. [15]

Eccles, J. and A. Wigfield (2002), "Motivational beliefs, values, and goals", *Annual Review of Psychology*, Vol. 53, pp. 109-132, https://doi.org/10.1146/annurev.psych.53.100901.135153. [17]

Eklöf, H. (2010), "Skill and will: Test-taking motivation and assessment quality", *Assessment in Education: Principles, Policy and Practice*, Vol. 17/4, pp. 345-356, https://doi.org/10.1080/0969594X.2010.516569. [13]

Eklöf, H., B. Pavešič and L. Grønmo (2014), "A Cross-National Comparison of Reported Effort and Mathematics Performance in TIMSS Advanced", *Applied Measurement in Education*, Vol. 27/1, pp. 31-45, https://doi.org/10.1080/08957347.2013.853070. [10]

Finn, B. (2015), "Measuring Motivation in Low-Stakes Assessments", *ETS Research Report Series*, Vol. 2015/2, pp. 1-17, https://doi.org/10.1002/ets2.12067. [5]

Fredricks, J. and W. McColskey (2012), "The measurement of student engagement: A comparative analysis of various methods and student self-report instruments", *Handbook of Research on Student Engagement*, pp. 763-782, https://doi.org/10.1007/978-1-4614-2018-7_37. [18]

Fronczyk, K. (2014), "The identification of random or careless responding in questionnaires: The example of the NEO-FFI 22", *Roczniki Psychologiczne/Annals of Psychology*, Vol. 17/2, pp. 457-473. [30]

Gneezy, U. et al. (2019), "Measuring Success in Education: The Role of Effort on the Test Itself", *American Economic Review: Insights*, Vol. 1/3, pp. 291-308, https://doi.org/10.1257/aeri.20180633. [14]

Gobert, J., R. Baker and M. Wixon (2015), "Operationalizing and Detecting Disengagement Within Online Science Microworlds", *Educational Psychologist*, Vol. 50/1, pp. 43-57, https://doi.org/10.1080/00461520.2014.999919. [29]

Hartig, J. and J. Buchholz (2012), "A multilevel item response model for item position effects and individual persistence", Vol. 54/4, pp. 418-431. [26]

Hitt, C. (n.d.), *Just Filling in the Bubbles: Using Careless Answer Patterns on Surveys as a Proxy Measure of Noncognitive Skills*, http://www.uaedreform.org/downloads/2015/07/edre-working-paper-2015-06.pdf. [41]

Hitt, C., J. Trivitt and A. Cheng (2016), "When you say nothing at all: The predictive power of student effort on surveys", *Economics of Education Review*, Vol. 52, pp. 105-119, https://doi.org/10.1016/J.ECONEDUREV.2016.02.001. [39]

Huang, J., M. Liu and N. Bowling (2015), "Insufficient effort responding: Examining an insidious confound in survey data", *Journal of Applied Psychology*, Vol. 100/3, pp. 828-845, https://doi.org/10.1037/a0038510. [16]

Johnson, T. et al. (2005), "The Relation Between Culture and Response Styles", *Journal of Cross-Cultural Psychology*, Vol. 36/2, pp. 264-277, https://doi.org/10.1177/0022022104272905. [57]

Kankaraš, M. (2017), "Personality matters: Relevance and assessment of personality characteristics", *OECD Education Working Papers*, No. 157, OECD Publishing, Paris, https://dx.doi.org/10.1787/8a294376-en. [32]

Knowles, E. (1988), "Item Context Effects on Personality Scales: Measuring Changes the Measure", *Journal of Personality and Social Psychology*, Vol. 55/2, pp. 312-320, https://doi.org/10.1037/0022-3514.55.2.312. [42]

Krosnick, J. (1991), "Response strategies for coping with the cognitive demands of attitude measures in surveys", *Applied Cognitive Psychology*, Vol. 5/3, pp. 213-236, https://doi.org/10.1002/ACP.2350050305. [43]

Kuhfeld, M. and J. Soland (2020), "Using Assessment Metadata to Quantify the Impact of Test Disengagement on Estimates of Educational Effectiveness", *Journal of Research on Educational Effectiveness*, Vol. 13/1, pp. 147-175, https://doi.org/10.1080/19345747.2019.1636437. [6]

Kyllonen, P. and H. Kell (2018), "Ability Tests Measure Personality, Personality Tests Measure Ability: Disentangling Construct and Method in Evaluating the Relationship between Personality and Ability", *Journal of Intelligence 2018, Vol. 6, Page 32*, Vol. 6/3, p. 32, https://doi.org/10.3390/JINTELLIGENCE6030032. [31]

Lau, A. et al. (2009), "Proctors Matter: Strategies for Increasing Examinee Effort on General Education Program Assessments", *The Journal of General Education*, Vol. 58/3, pp. 196-217, https://doi.org/10.1353/JGE.0.0045. [37]

Maddox, B. et al. (2018), "Observing response processes with eye tracking in international large-scale assessments: evidence from the OECD PIAAC assessment", *European Journal of Psychology of Education*, Vol. 33/3, pp. 543-558, https://doi.org/10.1007/S10212-018-0380-2/FIGURES/2. [38]

Maniaci, M. and R. Rogge (2014), "Caring about carelessness: Participant inattention and its effects on research", *Journal of Research in Personality*, Vol. 48/1, pp. 61-83, https://doi.org/10.1016/j.jrp.2013.09.008. [21]

Ocumpaugh, J. et al. (2015), *Baker Rodrigo Ocumpaugh Monitoring Protocol (BROMP) 2.0 Technical and Training Manual*, Technical Report. New York, NY: Teachers College, Columbia University. Manila, Philippines: Ateneo Laboratory for the Learning Sciences. [36]

OECD (2014), *PISA 2012 Technical Report*, https://www.oecd.org/pisa/pisaproducts/PISA-2012-technical-report-final.pdf (accessed on 25 March 2022). [35]

OECD (n.d.), *PISA 2018 Technical Report*, https://www.oecd.org/pisa/data/pisa2018technicalreport. [34]

Pintrich, P. and E. De Groot (1990), "Motivational and Self-Regulated Learning Components of Classroom Academic Performance", *Applied Measurement in Education*, Vol. 82/1, pp. 33-40, https://doi.org/10.1037/0022-0663.82.1.33. [7]

Pohl, S., E. Ulitzsch and M. von Davier (2021), *Reframing rankings in educational assessments*, American Association for the Advancement of Science, https://doi.org/10.1126/science.abd3300. [12]

R Core Team (2021), *R: A Language and Environment for Statistical Computing*, https://www.R-project.org/. [54]

Rios, J. et al. (2017), "Evaluating the Impact of Careless Responding on Aggregated-Scores: To Filter Unmotivated Examinees or Not?", *International Journal of Testing*, Vol. 17/1, pp. 74-104, https://doi.org/10.1080/15305058.2016.1231193. [47]

Rios, J. et al. (2017), "Evaluating the Impact of Careless Responding on Aggregated-Scores: To Filter Unmotivated Examinees or Not?", *International Journal of Testing*, Vol. 17/1, pp. 74-104, https://doi.org/10.1080/15305058.2016.1231193. [48]

Soland, J., S. Wise and L. Gao (2019), "Identifying Disengaged Survey Responses: New Evidence Using Response Time Metadata", *Applied Measurement in Education*, Vol. 32/2, pp. 151-165, https://doi.org/10.1080/08957347.2019.1577244. [2]

StataCorp (2021), *Stata Statistical Software: Release 17*, StataCorp LLC, College Station, TX. [53]

Ulitzsch, E., M. Davier and S. Pohl (2019), "A hierarchical latent response model for inferences about examinee engagement in terms of guessing and item-level non-response", *British Journal of Mathematical and Statistical Psychology*, Vol. 73/S1, pp. 83-112, https://doi.org/10.1111/bmsp.12188. [51]

Van de Vijver, F. and J. He (2014), "Report on Social Desirability, Midpoint and Extreme Responding in TALIS 2013", *OECD Education Working Papers*, No. 107, OECD Publishing, Paris, https://dx.doi.org/10.1787/5jxswcfwt76h-en. [55]

Van Der Linden, W. (2009), *Conceptual Issues in Response-Time Modeling*. [49]

Vannette, D. and J. Krosnick (2014), "A comparison of survey satisficing and mindlessness", *The Wiley Blackwell handbook of mindfulness*, Vol. 1, p. 312. [33]

Wise, S. (2019), "An Information-Based Approach to Identifying Rapid-Guessing Thresholds", *Applied Measurement in Education*, Vol. 32/4, pp. 325-336, https://doi.org/10.1080/08957347.2019.1660350. [52]

Wise, S. (2017), "Rapid-Guessing Behavior: Its Identification, Interpretation, and Implications", *Educational Measurement: Issues and Practice*, Vol. 36/4, pp. 52-61, https://doi.org/10.1111/emip.12165. [28]

Wise, S. and C. DeMars (2006), "An Application of Item Response Time: The Effort-Moderated IRT Model", *Journal of Educational Measurement,*, Vol. 34/1, pp. 19-38, https://doi.org/10.1111/j.1745-3984.2006.00002.x. [50]

Wise, S. and C. Demars (2005), "Low Examinee Effort in Low-Stakes Assessment: Problems and Potential Solutions", *Educational Assessment*, Vol. 10/1, pp. 1-17, https://doi.org/10.1207/s15326977ea1001_1. [8]

Wise, S., M. Kuhfeld and J. Soland (2019), "The Effects of Effort Monitoring With Proctor Notification on Test-Taking Engagement, Test Performance, and Validity", *Applied Measurement in Education*, Vol. 32/2, pp. 183-192, https://doi.org/10.1080/08957347.2019.1577248. [22]

Wolf, L. and J. Smith (1995), "The Consequence of Consequence: Motivation, Anxiety, and Test Performance", *Applied Measurement in Education*, Vol. 8/3, pp. 227-242, https://doi.org/10.1207/s15324818ame0803_3. [4]

Wu, Q. et al. (2019), "Predictors of individual performance changes related to item positions in PISA assessments", *Large-Scale Assessments in Education*, Vol. 7/1, pp. 1-21, https://doi.org/10.1186/s40536-019-0073-6. [45]

Zamarro, G., C. Hitt and I. Mendez (2019), "When Students Don't Care: Reexamining International Differences in Achievement and Student Effort", *Journal of Human Capital*, Vol. 13/4, pp. 519-552, https://doi.org/10.1086/705799. [9]

Annex A. Variation between countries/economies

Table A.1. Descriptive statistics for measures of disengagement in PISA 2018 (OECD countries)

Country / Economy	Self-report				Questionnaire-based behaviour						Test-based behaviour			
	Effort (R)		Work mastery (R)		Non-response		Non-differentiation		Inconsistency		Performance decline		Rapid guessing	
	Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.
Australia	**0.12**	0.01	**0.03**	0.01	**-0.04**	0.01	**0.27**	0.01	**-0.32**	0.01	**0.07**	0.01	0.01	0.01
Austria	**0.27**	0.02	**0.04**	0.01	-0.01	0.02	**-0.37**	0.02	**0.32**	0.01	**-0.15**	0.01	-0.03	0.02
Belgium	**0.19**	0.01	-	-	-0.03	0.02	**-0.14**	0.01	**0.11**	0.01	**-0.07**	0.01	**-0.11**	0.02
Canada	**0.10**	0.01	**-0.11**	0.01	**-0.05**	0.02	**0.04**	0.02	**-0.26**	0.01	**0.05**	0.01	**-0.12**	0.01
Chile	**-0.19**	0.02	**-0.29**	0.02	**0.17**	0.04	**0.10**	0.02	**0.05**	0.02	**0.11**	0.01	**0.12**	0.02
Colombia	**-0.44**	0.02	**-0.13**	0.02	**0.49**	0.08	**0.16**	0.02	**-0.10**	0.02	**0.14**	0.01	**0.22**	0.05
Czech Republic	**0.20**	0.02	**0.25**	0.01	-0.04	0.03	**-0.13**	0.02	0.01	0.02	**0.13**	0.02	-0.01	0.02
Denmark	**0.08**	0.02	**0.05**	0.01	**-0.13**	0.02	**0.05**	0.02	**-0.21**	0.01	-0.02	0.01	**-0.10**	0.02
Estonia	**-0.04**	0.02	**0.31**	0.01	**-0.18**	0.01	**-0.03**	0.01	**-0.11**	0.01	**-0.08**	0.01	**-0.18**	0.01
Finland	**-0.18**	0.02	**0.32**	0.02	**-0.08**	0.02	**0.19**	0.02	**-0.20**	0.01	**-0.16**	0.01	**-0.14**	0.01
France	**0.26**	0.02	**0.24**	0.01	**0.20**	0.03	**-0.24**	0.02	**0.40**	0.02	-0.02	0.01	0.03	0.02
Germany	**0.25**	0.02	**0.08**	0.02	**0.23**	0.03	**-0.50**	0.02	**0.31**	0.02	**-0.18**	0.02	-0.03	0.02
Greece	**0.08**	0.02	**-0.27**	0.02	**0.10**	0.04	**-0.26**	0.02	**0.27**	0.02	**0.12**	0.01	**0.29**	0.04
Hungary	-0.03	0.02	-0.02	0.01	**-0.08**	0.02	**-0.12**	0.02	**0.07**	0.02	**-0.12**	0.01	-0.04	0.02
Iceland	-0.01	0.02	**0.14**	0.02	0.03	0.02	**0.49**	0.02	**-0.33**	0.02	**0.09**	0.02	**0.07**	0.02
Ireland	**-0.18**	0.01	**0.09**	0.01	**-0.10**	0.01	**-0.05**	0.01	**-0.18**	0.01	**-0.18**	0.01	**-0.19**	0.01
Israel	**-0.12**	0.02	**-0.34**	0.02	**0.47**	0.06	**-0.07**	0.02	**0.28**	0.02	**0.34**	0.01	**0.41**	0.05
Italy	**-0.16**	0.02	**-0.48**	0.02	**0.23**	0.04	**-0.28**	0.02	**0.32**	0.02	0.00	0.01	**0.15**	0.04
Japan	**0.27**	0.02	**0.12**	0.02	**-0.21**	0.01	**0.11**	0.02	**0.18**	0.02	**-0.16**	0.01	**-0.08**	0.02
Korea	**-0.33**	0.02	**-0.39**	0.01	**-0.29**	0.01	**0.51**	0.02	**-0.57**	0.01	-0.03	0.02	**0.07**	0.03
Latvia	**-0.04**	0.02	**0.09**	0.01	**-0.10**	0.02	**-0.23**	0.02	**0.18**	0.02	**0.05**	0.01	**-0.08**	0.02
Lithuania	**-0.18**	0.01	0.02	0.01	**-0.07**	0.02	-0.01	0.02	**0.24**	0.02	**-0.06**	0.01	**-0.09**	0.02
Luxembourg	**0.36**	0.01	**0.26**	0.02	0.00	0.01	**-0.07**	0.01	**0.29**	0.01	**-0.06**	0.02	-0.01	0.01
Mexico	**-0.53**	0.01	**-0.37**	0.02	0.05	0.04	**0.15**	0.02	**-0.11**	0.02	**0.02**	0.01	**-0.20**	0.02
Netherlands	**0.13**	0.02	**0.40**	0.02	**-0.23**	0.01	0.01	0.02	**-0.15**	0.02	**0.04**	0.02	-0.03	0.03
New Zealand	**0.05**	0.02	**0.04**	0.01	**-0.09**	0.01	**0.06**	0.02	**-0.26**	0.01	0.00	0.01	-0.01	0.02
Norway	**0.14**	0.02	-0.01	0.01	0.03	0.03	**0.24**	0.02	**-0.04**	0.02	**0.11**	0.02	**0.14**	0.03
Poland	**0.11**	0.02	**-0.10**	0.02	**-0.16**	0.02	**-0.09**	0.02	0.00	0.02	**-0.10**	0.02	**-0.07**	0.02
Portugal	**0.08**	0.02	**-0.08**	0.01	**-0.11**	0.02	**0.14**	0.02	**-0.22**	0.02	**-0.06**	0.01	-0.03	0.02
Slovak Republic	**0.19**	0.02	**0.19**	0.01	**0.12**	0.03	**-0.05**	0.02	**0.17**	0.02	**0.14**	0.01	**0.14**	0.02
Slovenia	**0.05**	0.02	**-0.41**	0.01	**-0.17**	0.01	**0.03**	0.01	**-0.10**	0.01	0.02	0.02	**-0.14**	0.01
Spain	**0.05**	0.01	**-0.17**	0.01	**0.20**	0.04	**-0.18**	0.01	**0.10**	0.01	**0.13**	0.01	-0.02	0.01
Sweden	**0.13**	0.02	**0.30**	0.02	0.02	0.03	**0.13**	0.02	**-0.09**	0.02	0.01	0.01	**0.21**	0.04
Switzerland	**0.22**	0.02	**0.21**	0.02	**0.13**	0.03	**-0.34**	0.02	**0.25**	0.02	**0.07**	0.02	0.03	0.03
Turkey	**-0.68**	0.02	-0.02	0.02	**-0.21**	0.01	**0.11**	0.02	0.04	0.02	**-0.11**	0.01	**-0.22**	0.01
United Kingdom	**0.10**	0.02	**0.17**	0.01	**0.06**	0.02	**0.10**	0.01	**-0.14**	0.01	-0.02	0.01	**0.12**	0.02
United States	**-0.32**	0.01	**-0.18**	0.02	**-0.17**	0.02	**0.26**	0.02	**-0.22**	0.02	0.04	0.02	**-0.11**	0.02

Note: Question ST182 underlying the computation of Work Mastery has not been administered in Belgium.

Table A.2. Descriptive statistics for measures of disengagement in PISA 2018 (Partner countries and economies)

Country / Economy	Self-report				Questionnaire-based behaviour						Test-based behaviour			
	Effort (R)		Work mastery (R)		Non-response		Non-dif-ferentiation		Incon-sistency		Performance decline		Rapid guessing	
	Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.
Albania	**-0.72**	0.02	**-0.61**	0.02	**-0.07**	0.02	**0.35**	0.02	**-0.21**	0.02	**0.07**	0.01	**-0.04**	0.02
Argentina	**-0.11**	0.02	**-0.29**	0.02	**1.04**	0.06	**-0.30**	0.02	**0.47**	0.02	-	-	-	-
Baku (Azerbaijan)	**-0.21**	0.02	**-0.10**	0.02	**1.68**	0.07	**0.66**	0.02	0.04	0.02	**0.09**	0.02	**0.48**	0.03
Bosnia and Herzegovina	**0.14**	0.02	**-0.11**	0.02	**0.36**	0.03	**0.40**	0.02	**-0.15**	0.02	**-0.07**	0.01	**0.33**	0.03
Brazil	**-0.28**	0.02	**-0.26**	0.01	**1.02**	0.05	**0.18**	0.02	**0.04**	0.02	**0.11**	0.01	**0.33**	0.03
Brunei Darussalam	**0.13**	0.01	**-0.12**	0.01	**-0.06**	0.01	**-0.25**	0.01	**0.16**	0.01	**-0.26**	0.01	**-0.17**	0.01
B-S-J-Z (China)	**-0.72**	0.02	**-0.27**	0.02	**-0.35**	0.00	**0.30**	0.02	**-0.42**	0.02	**-0.25**	0.02	**-0.28**	0.01
Bulgaria	-0.04	0.02	**-0.06**	0.03	**0.89**	0.07	**0.37**	0.02	**0.27**	0.03	**-0.05**	0.01	**0.65**	0.06
Costa Rica	**-0.26**	0.01	**-0.60**	0.02	-0.04	0.03	**0.17**	0.02	**0.04**	0.02	**0.13**	0.00	**-0.11**	0.03
Croatia	0.02	0.02	**-0.21**	0.01	0.02	0.02	**0.17**	0.02	**-0.23**	0.01	**-0.13**	0.01	**-0.09**	0.02
Dominican Republic	**-0.68**	0.02	**-0.11**	0.03	**1.75**	0.11	**0.29**	0.02	**0.27**	0.02	**-0.14**	0.01	**1.10**	0.06
Georgia	**-0.05**	0.02	**-0.08**	0.02	**0.61**	0.04	**0.31**	0.03	**0.13**	0.03	**0.04**	0.01	**0.84**	0.05
Hong Kong (China)	**0.13**	0.02	**0.03**	0.01	**-0.30**	0.01	**0.46**	0.02	**-0.47**	0.02	**-0.08**	0.02	-0.04	0.02
Indonesia	**-0.60**	0.03	**-0.25**	0.02	-0.05	0.05	**0.32**	0.03	**-0.25**	0.03	**0.06**	0.00	**-0.09**	0.04
Jordan	**-0.38**	0.03	**-0.35**	0.03	**0.08**	0.03	-0.03	0.02	**0.57**	0.02	-	-	-	-
Kazakhstan	**-0.54**	0.01	**-0.03**	0.01	**0.11**	0.02	**0.47**	0.01	**-0.25**	0.01	**-0.17**	0.01	**-0.06**	0.01
Kosovo[7]	**-0.54**	0.02	**-0.36**	0.02	**0.25**	0.03	**0.21**	0.02	**-0.04**	0.02	**-0.20**	0.01	**0.05**	0.02
Lebanon	**0.42**	0.03	-0.04	0.03	**0.60**	0.05	**-0.59**	0.03	**1.43**	0.04	-	-	-	-
Macao (China)	**-0.25**	0.01	0.00	0.01	**-0.34**	0.00	**-0.14**	0.02	**-0.14**	0.01	**-0.22**	0.02	**-0.28**	0.01
Malaysia	**-0.50**	0.02	**-0.42**	0.02	**-0.27**	0.02	**-0.15**	0.02	0.01	0.02	**-0.20**	0.01	**-0.19**	0.01
Malta	**0.22**	0.02	**-0.29**	0.02	**0.30**	0.02	**0.14**	0.02	**-0.05**	0.02	**0.09**	0.01	**0.52**	0.03
Moldova	**-0.42**	0.01	**0.04**	0.01	**0.06**	0.02	**-0.19**	0.02	**0.06**	0.01	-	-	-	-
Montenegro	**-0.05**	0.02	**0.08**	0.01	**0.42**	0.02	**0.31**	0.01	-0.01	0.02	**-0.06**	0.01	**0.35**	0.02
Morocco	**-0.30**	0.02	**-0.36**	0.02	**1.59**	0.08	**-0.36**	0.02	**0.66**	0.03	**-0.19**	0.01	-0.03	0.02
North Macedonia	**-0.21**	0.02	**-0.51**	0.01	**0.17**	0.02	**0.10**	0.02	**0.54**	0.02	-	-	-	-
Panama	**-0.38**	0.01	**-0.40**	0.02	**1.07**	0.05	**-0.08**	0.02	**0.35**	0.02	**-0.12**	0.01	**0.16**	0.04
Peru	**-0.29**	0.01	**-0.39**	0.02	-0.01	0.02	0.03	0.02	**-0.11**	0.01	0.01	0.01	**-0.16**	0.01
Philippines	**-0.61**	0.02	**-0.07**	0.02	**0.06**	0.03	**0.25**	0.02	**-0.16**	0.03	**-0.27**	0.01	-0.03	0.02
Qatar	**-0.16**	0.01	**-0.28**	0.01	**0.44**	0.01	**0.19**	0.01	**0.44**	0.01	**0.16**	0.01	**0.91**	0.01
Romania	**-0.12**	0.02	0.01	0.02	**0.10**	0.03	**-0.20**	0.02	**0.08**	0.03	-	-	-	-
Saudi Arabia	**-0.40**	0.02	**-0.22**	0.02	**0.28**	0.05	**-0.32**	0.02	**0.77**	0.02	-	-	-	-
Serbia	**0.12**	0.02	**-0.07**	0.02	**0.68**	0.07	**0.22**	0.02	**0.05**	0.02	**0.22**	0.01	**0.39**	0.05
Singapore	**0.06**	0.01	**-0.23**	0.01	**-0.29**	0.01	**0.11**	0.01	**-0.36**	0.01	**-0.17**	0.01	**-0.14**	0.01
Chinese Taipei	**-0.35**	0.02	**-0.19**	0.01	**-0.30**	0.01	**0.46**	0.02	**-0.56**	0.01	**-0.17**	0.02	-0.02	0.02
Thailand	**-0.53**	0.02	**-0.29**	0.02	**-0.26**	0.02	**0.59**	0.03	**-0.46**	0.02	**0.12**	0.01	-0.03	0.03
Ukraine	**-0.23**	0.02	**0.04**	0.02	**0.15**	0.03	**-0.26**	0.02	**0.22**	0.02	-	-	-	-
United Arab Emirates	**-0.46**	0.05	**-0.40**	0.01	**0.11**	0.02	**0.24**	0.01	**0.14**	0.01	0.03	0.02	**0.22**	0.02
Uruguay	**-0.12**	0.02	**-0.27**	0.02	**0.93**	0.06	**0.11**	0.02	**0.17**	0.02	0.01	0.01	**0.20**	0.03
Viet Nam	**-0.70**	0.02	**-0.18**	0.02	**-0.31**	0.01	**-0.59**	0.02	**0.19**	0.02	-	-	-	-

Note: Performance decline and Rapid guessing could only be computed for countries/economies with a computer-based assessment.

[7] This designation is without prejudice to positions on status, and is in line with UNSCR 1244/1999 and the ICJ Opinion on the Kosovo declaration of independence.

Annex B. Differences in disengagement by gender

Table B.1. Gender disparities (female – male) in PISA 2018 for different measures of disengagement (OECD countries)

Country / Economy	Self-report				Questionnaire-based behaviour						Test-based behaviour			
	Effort (R)		Work mastery (R)		Non-response		Non-dif-ferentiation		Incon-sistency		Performance decline		Rapid guessing	
	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.
Australia	**-0.05**	(0.02)	**-0.15**	(0.02)	**-0.09**	(0.02)	**-0.09**	(0.02)	**-0.16**	(0.02)	**-0.06**	(0.02)	**-0.18**	(0.02)
Austria	**-0.19**	(0.04)	**-0.08**	(0.03)	**-0.24**	(0.04)	**-0.11**	(0.03)	**-0.24**	(0.03)	-0.03	(0.03)	**-0.19**	(0.03)
Belgium	**-0.11**	(0.02)	-	-	**-0.12**	(0.03)	0.00	(0.02)	**-0.23**	(0.02)	-0.04	(0.03)	**-0.12**	(0.03)
Canada	**-0.15**	(0.02)	**-0.16**	(0.02)	**-0.17**	(0.02)	-0.01	(0.02)	**-0.22**	(0.02)	**-0.04**	(0.02)	**-0.12**	(0.01)
Chile	**-0.14**	(0.03)	**-0.18**	(0.03)	**-0.10**	(0.03)	0.02	(0.03)	**-0.22**	(0.03)	-0.01	(0.01)	-0.03	(0.03)
Colombia	**-0.13**	(0.02)	**-0.14**	(0.03)	-0.05	(0.05)	-0.02	(0.04)	**-0.22**	(0.03)	-0.02	(0.01)	-0.03	(0.04)
Czech Republic	**-0.12**	(0.03)	**-0.10**	(0.03)	**-0.09**	(0.04)	**-0.11**	(0.03)	**-0.38**	(0.03)	**-0.14**	(0.03)	**-0.13**	(0.03)
Denmark	**-0.12**	(0.03)	**-0.09**	(0.03)	**-0.14**	(0.02)	**-0.19**	(0.02)	**-0.11**	(0.03)	**-0.11**	(0.03)	**-0.14**	(0.02)
Estonia	**-0.20**	(0.03)	**-0.15**	(0.03)	**-0.10**	(0.02)	**-0.12**	(0.03)	**-0.21**	(0.03)	-0.03	(0.03)	**-0.06**	(0.02)
Finland	**-0.23**	(0.03)	**-0.18**	(0.02)	**-0.18**	(0.03)	**-0.36**	(0.03)	**-0.07**	(0.02)	**-0.14**	(0.03)	**-0.21**	(0.03)
France	**-0.24**	(0.03)	**-0.18**	(0.02)	**-0.27**	(0.04)	**-0.08**	(0.03)	**-0.21**	(0.02)	**-0.06**	(0.03)	**-0.15**	(0.03)
Germany	**-0.28**	(0.03)	**-0.09**	(0.03)	**-0.24**	(0.04)	**-0.11**	(0.03)	**-0.19**	(0.03)	**-0.06**	(0.03)	**-0.12**	(0.03)
Greece	**-0.12**	(0.02)	**-0.20**	(0.02)	**-0.28**	(0.04)	-0.02	(0.03)	**-0.35**	(0.04)	-0.01	(0.02)	**-0.27**	(0.05)
Hungary	**-0.15**	(0.03)	0.04	(0.03)	**-0.12**	(0.03)	**-0.09**	(0.03)	**-0.17**	(0.03)	**-0.10**	(0.03)	**-0.15**	(0.03)
Iceland	**-0.15**	(0.04)	-0.04	(0.04)	**-0.24**	(0.04)	**-0.27**	(0.03)	**-0.10**	(0.04)	-0.06	(0.04)	**-0.24**	(0.04)
Ireland	**-0.17**	(0.03)	**-0.11**	(0.03)	**-0.11**	(0.02)	-0.04	(0.03)	**-0.17**	(0.02)	-0.03	(0.03)	**-0.12**	(0.02)
Israel	**-0.19**	(0.03)	**-0.30**	(0.03)	**-0.66**	(0.07)	**-0.12**	(0.04)	**-0.32**	(0.03)	0.02	(0.02)	**-0.70**	(0.08)
Italy	**-0.09**	(0.03)	**-0.26**	(0.03)	**-0.10**	(0.05)	0.01	(0.03)	**-0.38**	(0.04)	**-0.07**	(0.03)	**-0.21**	(0.04)
Japan	**-0.09**	(0.03)	**-0.14**	(0.03)	**-0.05**	(0.02)	**-0.22**	(0.03)	**-0.19**	(0.02)	**-0.07**	(0.03)	**-0.13**	(0.03)
Korea	0.06	(0.03)	**0.06**	(0.03)	**-0.04**	(0.01)	**-0.31**	(0.04)	**-0.06**	(0.02)	**-0.11**	(0.03)	**-0.22**	(0.06)
Latvia	**-0.18**	(0.03)	**-0.15**	(0.03)	**-0.13**	(0.03)	**-0.11**	(0.03)	**-0.28**	(0.03)	-0.01	(0.03)	**-0.15**	(0.03)
Lithuania	**-0.21**	(0.02)	**-0.13**	(0.02)	**-0.18**	(0.02)	**-0.12**	(0.03)	**-0.29**	(0.03)	**-0.05**	(0.02)	**-0.24**	(0.02)
Luxembourg	**-0.21**	(0.03)	**-0.21**	(0.03)	**-0.11**	(0.02)	**-0.10**	(0.03)	**-0.35**	(0.03)	**-0.09**	(0.03)	**-0.25**	(0.02)
Mexico	**-0.10**	(0.02)	**-0.24**	(0.03)	-0.02	(0.03)	**0.08**	(0.03)	**-0.20**	(0.04)	-0.01	(0.02)	-0.01	(0.02)
Netherlands	**-0.12**	(0.03)	-0.03	(0.03)	**-0.07**	(0.02)	**-0.08**	(0.02)	**-0.18**	(0.03)	**-0.07**	(0.03)	**-0.20**	(0.04)
New Zealand	**-0.07**	(0.03)	**-0.19**	(0.03)	**-0.12**	(0.02)	-0.04	(0.03)	**-0.17**	(0.03)	**-0.08**	(0.02)	**-0.19**	(0.04)
Norway	**-0.17**	(0.03)	**-0.12**	(0.03)	**-0.26**	(0.04)	**-0.09**	(0.02)	**-0.22**	(0.03)	**-0.07**	(0.03)	**-0.26**	(0.04)
Poland	**-0.22**	(0.03)	**-0.22**	(0.03)	**-0.14**	(0.03)	0.00	(0.03)	**-0.40**	(0.03)	-0.07	(0.04)	**-0.19**	(0.03)
Portugal	**-0.27**	(0.03)	**-0.14**	(0.03)	**-0.09**	(0.02)	**-0.07**	(0.03)	**-0.12**	(0.03)	-0.06	(0.03)	**-0.16**	(0.02)
Slovak Republic	**-0.15**	(0.03)	**-0.10**	(0.03)	**-0.14**	(0.03)	**-0.07**	(0.03)	**-0.37**	(0.04)	-0.03	(0.02)	**-0.16**	(0.03)
Slovenia	-0.04	(0.03)	**-0.25**	(0.03)	**-0.12**	(0.01)	**-0.14**	(0.02)	**-0.28**	(0.03)	**-0.14**	(0.03)	**-0.19**	(0.02)
Spain	**-0.13**	(0.02)	**-0.26**	(0.02)	**-0.08**	(0.03)	**-0.03**	(0.02)	**-0.22**	(0.02)	**-0.03**	(0.01)	**-0.15**	(0.02)
Sweden	**-0.15**	(0.04)	0.00	(0.03)	**-0.22**	(0.04)	**-0.25**	(0.03)	**-0.17**	(0.02)	**-0.07**	(0.03)	**-0.14**	(0.03)
Switzerland	**-0.13**	(0.04)	**-0.14**	(0.03)	**-0.16**	(0.03)	**-0.11**	(0.03)	**-0.25**	(0.03)	**-0.08**	(0.03)	**-0.20**	(0.03)
Turkey	**-0.11**	(0.03)	**-0.30**	(0.03)	**-0.11**	(0.02)	-0.05	(0.03)	**-0.34**	(0.03)	**-0.10**	(0.03)	**-0.09**	(0.02)
United Kingdom	**-0.16**	(0.02)	**-0.09**	(0.02)	**-0.16**	(0.04)	**-0.05**	(0.03)	**-0.14**	(0.02)	**-0.07**	(0.03)	**-0.19**	(0.04)
United States	**-0.07**	(0.03)	**-0.12**	(0.03)	**-0.11**	(0.02)	-0.02	(0.03)	**-0.15**	(0.03)	-0.04	(0.04)	**-0.19**	(0.03)
OECD average	**-0.14**	(0.00)	**-0.14**	(0.00)	**-0.15**	(0.01)	**-0.09**	(0.00)	**-0.22**	(0.00)	**-0.06**	(0.00)	**-0.18**	(0.01)

Note: Question ST182 underlying the computation of Work Mastery has not been administered in Belgium.

Table B.2. Gender disparities (female – male) in PISA 2018 for different measures of disengagement (Partner countries and economies)

Country / Economy	Self-report				Questionnaire-based behaviour						Test-based behaviour			
	Effort (R)		Work mastery (R)		Non-response		Non-dif-ferentiation		Incon-sistency		Performance decline		Rapid guessing	
	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.
Albania	**-0.17**	(0.03)	**-0.35**	(0.03)	**-0.25**	(0.03)	**-0.21**	(0.03)	**-0.24**	(0.04)	-0.01	(0.02)	**-0.33**	(0.03)
Argentina	**-0.17**	(0.03)	**-0.19**	(0.03)	-0.09	(0.06)	-0.01	(0.02)	**-0.25**	(0.03)	-	-	-	-
Baku (Azerbaijan)	**-0.08**	(0.03)	**-0.10**	(0.04)	**-0.61**	(0.06)	**-0.27**	(0.03)	**-0.36**	(0.03)	**-0.10**	(0.03)	**-0.52**	(0.03)
Bosnia and Herzegovina	**-0.25**	(0.04)	**-0.20**	(0.03)	**-0.32**	(0.04)	-0.05	(0.03)	**-0.38**	(0.03)	-0.03	(0.02)	**-0.29**	(0.03)
Brazil	**-0.12**	(0.02)	**-0.18**	(0.03)	**-0.31**	(0.05)	0.00	(0.03)	**-0.26**	(0.03)	**-0.05**	(0.02)	**-0.17**	(0.03)
Brunei Darussalam	**-0.07**	(0.02)	**-0.21**	(0.02)	**-0.14**	(0.02)	-0.04	(0.02)	**-0.12**	(0.02)	-0.02	(0.01)	**-0.15**	(0.02)
B-S-J-Z (China)	**-0.07**	(0.02)	0.01	(0.03)	**-0.01**	(0.00)	**-0.08**	(0.02)	**-0.19**	(0.03)	-0.06	(0.03)	**-0.04**	(0.01)
Bulgaria	**-0.23**	(0.04)	**-0.32**	(0.04)	**-0.43**	(0.08)	-0.02	(0.04)	**-0.55**	(0.04)	**-0.08**	(0.03)	**-0.52**	(0.06)
Costa Rica	**-0.17**	(0.03)	**-0.19**	(0.03)	**-0.12**	(0.04)	-0.02	(0.03)	**-0.18**	(0.02)	0.00	(0.01)	**-0.07**	(0.02)
Croatia	**-0.13**	(0.03)	**-0.10**	(0.03)	**-0.23**	(0.04)	**-0.13**	(0.02)	**-0.23**	(0.02)	-0.02	(0.02)	**-0.23**	(0.03)
Dominican Republic	**-0.16**	(0.03)	**-0.24**	(0.05)	**-0.14**	(0.07)	-0.01	(0.05)	**-0.32**	(0.05)	0.02	(0.01)	-0.02	(0.05)
Georgia	**-0.17**	(0.04)	**-0.34**	(0.04)	**-0.65**	(0.06)	**-0.12**	(0.04)	**-0.46**	(0.04)	0.00	(0.03)	**-0.64**	(0.05)
Hong Kong (China)	-0.03	(0.03)	-0.04	(0.02)	**-0.08**	(0.01)	**-0.21**	(0.04)	**-0.23**	(0.03)	**-0.11**	(0.03)	**-0.28**	(0.03)
Indonesia	-0.04	(0.04)	**-0.21**	(0.03)	**-0.10**	(0.04)	**-0.13**	(0.04)	**-0.23**	(0.04)	0.02	(0.01)	**-0.19**	(0.03)
Jordan	**-0.21**	(0.06)	**-0.52**	(0.05)	**-0.23**	(0.06)	**-0.12**	(0.05)	**-0.60**	(0.05)	-	-	-	-
Kazakhstan	**-0.17**	(0.02)	**-0.24**	(0.02)	**-0.16**	(0.02)	**-0.22**	(0.02)	**-0.12**	(0.02)	0.00	(0.01)	**-0.14**	(0.02)
Kosovo	**-0.20**	(0.04)	**-0.35**	(0.03)	**-0.33**	(0.04)	**-0.23**	(0.04)	**-0.34**	(0.03)	0.01	(0.01)	**-0.16**	(0.03)
Lebanon	**-0.15**	(0.04)	**-0.19**	(0.05)	**0.23**	(0.06)	**0.09**	(0.04)	**-0.24**	(0.06)	-	-	-	-
Macao (China)	**-0.06**	(0.03)	**-0.15**	(0.02)	**-0.02**	(0.01)	-0.01	(0.03)	**-0.27**	(0.03)	-0.03	(0.05)	**-0.06**	(0.01)
Malaysia	-0.03	(0.03)	**-0.31**	(0.03)	-0.02	(0.01)	-0.05	(0.03)	**-0.10**	(0.03)	-0.04	(0.02)	**-0.10**	(0.02)
Malta	**-0.27**	(0.04)	**-0.21**	(0.03)	**-0.63**	(0.05)	-0.04	(0.04)	**-0.37**	(0.04)	-0.03	(0.03)	**-0.80**	(0.05)
Moldova	**-0.21**	(0.03)	**-0.13**	(0.03)	**-0.28**	(0.03)	-0.06	(0.03)	**-0.18**	(0.03)	-	-	-	-
Montenegro	**-0.16**	(0.03)	**-0.08**	(0.03)	**-0.47**	(0.04)	**-0.17**	(0.03)	**-0.37**	(0.03)	0.02	(0.02)	**-0.35**	(0.03)
Morocco	**-0.13**	(0.03)	**-0.29**	(0.03)	-0.07	(0.05)	**-0.06**	(0.03)	**-0.28**	(0.04)	**0.04**	(0.01)	**-0.11**	(0.02)
North Macedonia	**-0.26**	(0.03)	**-0.29**	(0.03)	**-0.15**	(0.03)	**-0.09**	(0.03)	**-0.25**	(0.03)	-	-	-	-
Panama	**-0.11**	(0.03)	**-0.22**	(0.04)	0.05	(0.06)	0.02	(0.04)	**-0.13**	(0.03)	0.01	(0.01)	-0.06	(0.05)
Peru	**-0.10**	(0.03)	**-0.17**	(0.03)	0.03	(0.02)	0.00	(0.03)	**-0.09**	(0.03)	-0.03	(0.03)	**-0.05**	(0.02)
Philippines	**-0.15**	(0.03)	**-0.23**	(0.02)	0.02	(0.02)	**0.07**	(0.04)	**-0.28**	(0.03)	**0.03**	(0.01)	**-0.14**	(0.02)
Qatar	**-0.16**	(0.02)	**-0.42**	(0.02)	**-0.77**	(0.03)	**-0.18**	(0.03)	**-0.51**	(0.02)	**-0.09**	(0.02)	**-1.15**	(0.03)
Romania	**-0.30**	(0.03)	**-0.13**	(0.03)	**-0.10**	(0.04)	**0.13**	(0.03)	**-0.39**	(0.03)	-	-	-	-
Saudi Arabia	**-0.23**	(0.04)	**-0.29**	(0.04)	**-0.36**	(0.09)	**-0.29**	(0.04)	**-0.41**	(0.04)	-	-	-	-
Serbia	**-0.12**	(0.04)	**-0.20**	(0.03)	**-0.47**	(0.07)	**-0.10**	(0.03)	**-0.40**	(0.04)	0.00	(0.01)	**-0.32**	(0.05)
Singapore	**-0.13**	(0.03)	**-0.16**	(0.02)	**-0.04**	(0.01)	-0.04	(0.02)	**-0.23**	(0.02)	-0.03	(0.03)	**-0.18**	(0.02)
Chinese Taipei	-0.01	(0.03)	-0.04	(0.02)	**-0.05**	(0.02)	**-0.18**	(0.04)	**-0.15**	(0.03)	-0.04	(0.03)	**-0.27**	(0.03)
Thailand	**-0.12**	(0.03)	**-0.30**	(0.03)	**-0.06**	(0.02)	**-0.15**	(0.04)	**-0.22**	(0.03)	-0.02	(0.02)	**-0.29**	(0.06)
Ukraine	**-0.35**	(0.03)	**-0.10**	(0.02)	**-0.20**	(0.03)	-0.01	(0.03)	**-0.31**	(0.03)	-	-	-	-
United Arab Emirates	**-0.28**	(0.09)	**-0.42**	(0.02)	**-0.55**	(0.03)	**-0.16**	(0.03)	**-0.45**	(0.02)	**-0.09**	(0.03)	**-0.63**	(0.03)
Uruguay	**-0.18**	(0.03)	**-0.14**	(0.03)	**-0.49**	(0.06)	-0.01	(0.03)	**-0.33**	(0.04)	-0.04	(0.02)	**-0.11**	(0.05)
Viet Nam	**-0.09**	(0.02)	**-0.10**	(0.02)	**-0.03**	(0.01)	**-0.09**	(0.03)	**-0.15**	(0.03)	-	-	-	-

Note: Performance decline and Rapid guessing could only be computed for countries/economies with a computer-based assessment.

Annex C. Differences in disengagement by socio-economic status (SES)

Table C.1. Differences between SES groups (top quarter – bottom quarter) in PISA 2018 for different measures of disengagement (OECD countries)

Country / Economy	Self-report				Questionnaire-based behaviour						Test-based behaviour			
	Effort (R)		Work mastery (R)		Non-response		Non-dif-ferentiation		Incon-sistency		Performance decline		Rapid guessing	
	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.
Australia	**-0.14**	(0.04)	**-0.37**	(0.04)	**-0.18**	(0.04)	**-0.15**	(0.05)	**-0.20**	(0.03)	**-0.23**	(0.04)	**-0.25**	(0.04)
Austria	**-0.16**	(0.05)	**-0.14**	(0.05)	**-0.15**	(0.05)	-0.03	(0.04)	**-0.16**	(0.05)	**-0.18**	(0.04)	**-0.23**	(0.06)
Belgium	**-0.12**	(0.03)	-	-	**-0.21**	(0.03)	**-0.07**	(0.03)	**-0.22**	(0.03)	**-0.33**	(0.04)	**-0.15**	(0.04)
Canada	**-0.11**	(0.04)	**-0.39**	(0.05)	**-0.31**	(0.06)	0.01	(0.05)	**-0.30**	(0.06)	**-0.08**	(0.03)	**-0.22**	(0.05)
Chile	**0.16**	(0.04)	**-0.23**	(0.04)	**-0.18**	(0.06)	**-0.18**	(0.04)	-0.02	(0.04)	**0.13**	(0.02)	**-0.16**	(0.05)
Colombia	0.07	(0.04)	**-0.28**	(0.07)	**-0.80**	(0.11)	**-0.16**	(0.06)	0.04	(0.05)	**0.14**	(0.02)	**-0.23**	(0.07)
Czech Republic	0.04	(0.05)	**-0.30**	(0.05)	**-0.15**	(0.06)	**-0.18**	(0.05)	**-0.27**	(0.05)	**-0.25**	(0.06)	**-0.31**	(0.07)
Denmark	**-0.31**	(0.09)	**-0.26**	(0.07)	**-0.25**	(0.05)	-0.09	(0.05)	**-0.24**	(0.07)	**-0.29**	(0.07)	**-0.34**	(0.09)
Estonia	**-0.19**	(0.06)	**-0.33**	(0.05)	**-0.10**	(0.04)	-0.03	(0.05)	**-0.15**	(0.06)	**-0.12**	(0.05)	-0.05	(0.04)
Finland	**-0.31**	(0.07)	**-0.55**	(0.06)	**-0.38**	(0.08)	**-0.15**	(0.07)	**-0.35**	(0.07)	**-0.16**	(0.06)	**-0.21**	(0.06)
France	**-0.21**	(0.05)	**-0.23**	(0.05)	**-0.53**	(0.07)	**-0.11**	(0.04)	**-0.46**	(0.05)	**-0.11**	(0.04)	**-0.32**	(0.06)
Germany	**-0.11**	(0.05)	**-0.12**	(0.04)	**-0.31**	(0.06)	-0.04	(0.04)	**-0.20**	(0.05)	**-0.10**	(0.04)	**-0.10**	(0.05)
Greece	0.03	(0.04)	**-0.25**	(0.04)	**-0.22**	(0.07)	0.02	(0.04)	**-0.19**	(0.04)	-0.03	(0.03)	**-0.28**	(0.06)
Hungary	-0.03	(0.05)	**-0.26**	(0.04)	**-0.15**	(0.06)	**-0.19**	(0.05)	**-0.11**	(0.05)	**-0.13**	(0.04)	**-0.32**	(0.07)
Iceland	**-0.24**	(0.11)	**-0.56**	(0.10)	-0.09	(0.07)	-0.09	(0.08)	**-0.29**	(0.10)	0.12	(0.07)	-0.19	(0.10)
Ireland	-0.06	(0.04)	**-0.28**	(0.05)	**-0.15**	(0.04)	-0.02	(0.04)	**-0.10**	(0.04)	**-0.17**	(0.04)	**-0.16**	(0.05)
Israel	**0.12**	(0.05)	**-0.33**	(0.07)	**-0.58**	(0.14)	**-0.25**	(0.07)	**-0.40**	(0.07)	**0.09**	(0.04)	**-0.57**	(0.13)
Italy	**-0.10**	(0.04)	**-0.12**	(0.04)	**-0.25**	(0.08)	-0.05	(0.05)	**-0.23**	(0.05)	**-0.12**	(0.04)	**-0.29**	(0.07)
Japan	0.04	(0.06)	**-0.23**	(0.05)	**-0.06**	(0.03)	-0.11	(0.06)	**-0.11**	(0.05)	**-0.12**	(0.05)	**-0.22**	(0.05)
Korea	**-0.12**	(0.05)	**-0.35**	(0.04)	0.01	(0.03)	-0.11	(0.07)	**-0.11**	(0.04)	**-0.29**	(0.05)	**-0.30**	(0.08)
Latvia	**-0.17**	(0.04)	**-0.30**	(0.05)	0.05	(0.04)	**0.12**	(0.05)	**-0.21**	(0.05)	-0.02	(0.05)	**-0.10**	(0.04)
Lithuania	**-0.17**	(0.04)	**-0.12**	(0.05)	-0.01	(0.05)	-0.01	(0.05)	**-0.13**	(0.05)	**0.06**	(0.03)	**-0.17**	(0.04)
Luxembourg	**-0.14**	(0.04)	**-0.22**	(0.04)	**-0.10**	(0.04)	-0.04	(0.04)	**-0.30**	(0.04)	**-0.28**	(0.05)	**-0.24**	(0.03)
Mexico	0.05	(0.03)	**-0.30**	(0.05)	**-0.23**	(0.07)	0.03	(0.06)	-0.09	(0.06)	**0.10**	(0.02)	-0.05	(0.03)
Netherlands	**-0.17**	(0.07)	0.07	(0.07)	**-0.14**	(0.06)	-0.13	(0.07)	**-0.16**	(0.07)	**-0.30**	(0.07)	**-0.36**	(0.11)
New Zealand	**-0.12**	(0.04)	**-0.39**	(0.04)	-0.07	(0.04)	**-0.14**	(0.04)	-0.06	(0.04)	**-0.17**	(0.04)	**-0.28**	(0.05)
Norway	**-0.28**	(0.10)	**-0.42**	(0.08)	**-0.46**	(0.09)	**-0.26**	(0.10)	**-0.39**	(0.09)	-0.09	(0.07)	**-0.38**	(0.09)
Poland	0.00	(0.05)	**-0.32**	(0.04)	**-0.09**	(0.04)	0.05	(0.04)	**-0.23**	(0.05)	-0.07	(0.05)	**-0.09**	(0.04)
Portugal	0.00	(0.04)	**-0.25**	(0.04)	**-0.11**	(0.03)	**-0.27**	(0.03)	**0.09**	(0.03)	**-0.12**	(0.04)	**-0.08**	(0.04)
Slovak Republic	0.00	(0.06)	**-0.27**	(0.04)	**-0.26**	(0.08)	0.02	(0.05)	**-0.35**	(0.05)	0.02	(0.04)	**-0.30**	(0.07)
Slovenia	**-0.18**	(0.06)	**-0.24**	(0.05)	**-0.11**	(0.04)	**-0.14**	(0.05)	**-0.13**	(0.06)	**-0.29**	(0.06)	**-0.21**	(0.04)
Spain	**0.06**	(0.03)	**-0.19**	(0.02)	**-0.29**	(0.07)	0.01	(0.03)	**-0.24**	(0.02)	0.02	(0.01)	**-0.10**	(0.03)
Sweden	**-0.21**	(0.08)	**-0.19**	(0.06)	**-0.41**	(0.09)	**-0.15**	(0.06)	**-0.24**	(0.06)	-0.05	(0.06)	**-0.13**	(0.06)
Switzerland	**-0.18**	(0.07)	-0.08	(0.06)	**-0.30**	(0.06)	**-0.10**	(0.05)	-0.07	(0.06)	**-0.26**	(0.05)	**-0.24**	(0.06)
Turkey	**0.30**	(0.06)	-0.09	(0.05)	-0.04	(0.03)	0.07	(0.07)	**-0.16**	(0.06)	-0.05	(0.07)	-0.05	(0.03)
United Kingdom	-0.08	(0.05)	**-0.25**	(0.04)	**-0.13**	(0.04)	-0.03	(0.04)	**-0.16**	(0.04)	**-0.19**	(0.04)	**-0.29**	(0.06)
United States	-0.01	(0.05)	**-0.33**	(0.05)	**-0.09**	(0.04)	-0.05	(0.04)	**-0.22**	(0.05)	**-0.43**	(0.06)	**-0.22**	(0.03)
OECD average	**-0.08**	(0.01)	**-0.26**	(0.01)	**-0.21**	(0.01)	**-0.08**	(0.01)	**-0.19**	(0.01)	**-0.12**	(0.01)	**-0.22**	(0.01)

Note: Question ST182 underlying the computation of Work Mastery has not been administered in Belgium.

Table C.2. Differences between SES groups (top quarter – bottom quarter) in PISA 2018 for different measures of disengagement (Partner countries and economies)

Country / Economy	Self-report				Questionnaire-based behaviour						Test-based behaviour			
	Effort (R)		Work mastery (R)		Non-response		Non-dif-ferentiation		Incon-sistency		Performance decline		Rapid guessing	
	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.
Albania	0.01	(0.06)	**-0.26**	(0.04)	0.02	(0.06)	**0.22**	(0.06)	**-0.13**	(0.06)	0.05	(0.03)	0.02	(0.05)
Argentina	-0.04	(0.04)	**-0.18**	(0.04)	**-0.74**	(0.11)	-0.06	(0.05)	**-0.16**	(0.05)	-	-	-	-
Baku (Azerbaijan)	**0.14**	(0.06)	**-0.13**	(0.06)	-0.04	(0.16)	**0.29**	(0.09)	**-0.18**	(0.07)	0.07	(0.07)	0.06	(0.07)
Bosnia and Herzegovina	0.09	(0.06)	**-0.17**	(0.05)	-0.17	(0.09)	-0.02	(0.06)	-0.09	(0.06)	**0.11**	(0.04)	**-0.12**	(0.06)
Brazil	0.05	(0.05)	**-0.25**	(0.05)	**-0.48**	(0.16)	**-0.17**	(0.05)	0.02	(0.04)	**0.11**	(0.04)	-0.03	(0.10)
Brunei Darussalam	**-0.19**	(0.03)	**-0.29**	(0.03)	**-0.19**	(0.03)	**-0.09**	(0.04)	**-0.07**	(0.04)	0.04	(0.02)	**-0.12**	(0.03)
B-S-J-Z (China)	**0.11**	(0.04)	**-0.31**	(0.03)	-0.01	(0.00)	-0.02	(0.03)	-0.02	(0.03)	**-0.26**	(0.06)	-0.03	(0.02)
Bulgaria	0.07	(0.06)	**-0.42**	(0.06)	-0.01	(0.13)	-0.08	(0.08)	**-0.27**	(0.08)	**-0.08**	(0.04)	**-0.52**	(0.08)
Costa Rica	**0.15**	(0.05)	0.02	(0.05)	**-0.21**	(0.05)	**-0.25**	(0.03)	0.00	(0.03)	**0.13**	(0.01)	**-0.19**	(0.05)
Croatia	-0.05	(0.05)	**-0.13**	(0.04)	**-0.15**	(0.04)	**-0.12**	(0.05)	-0.03	(0.04)	-0.06	(0.03)	-0.07	(0.04)
Dominican Republic	**0.17**	(0.06)	**-0.41**	(0.07)	**-0.92**	(0.17)	**-0.24**	(0.08)	-0.02	(0.09)	**0.12**	(0.03)	**-0.31**	(0.15)
Georgia	**0.27**	(0.07)	**-0.44**	(0.05)	-0.11	(0.08)	-0.04	(0.06)	-0.03	(0.07)	**0.14**	(0.04)	-0.09	(0.08)
Hong Kong (China)	0.07	(0.05)	**-0.14**	(0.04)	0.02	(0.02)	**-0.13**	(0.06)	0.02	(0.04)	**-0.16**	(0.05)	**-0.13**	(0.06)
Indonesia	0.22	(0.14)	**-0.32**	(0.08)	0.74	(0.41)	-0.11	(0.15)	0.17	(0.11)	**0.08**	(0.02)	0.07	(0.11)
Jordan	0.06	(0.07)	**-0.36**	(0.06)	0.04	(0.08)	0.05	(0.07)	-0.12	(0.06)	-	-	-	-
Kazakhstan	-0.03	(0.03)	**-0.31**	(0.04)	**-0.18**	(0.04)	0.06	(0.04)	-0.03	(0.04)	**0.07**	(0.02)	0.05	(0.03)
Kosovo	-0.08	(0.05)	**-0.40**	(0.06)	-0.10	(0.07)	**0.25**	(0.06)	**-0.16**	(0.06)	**0.08**	(0.03)	-0.04	(0.04)
Lebanon	**-0.27**	(0.08)	**-0.30**	(0.07)	**-0.40**	(0.11)	0.10	(0.07)	**-0.59**	(0.09)	-	-	-	-
Macao (China)	-0.05	(0.05)	**-0.27**	(0.04)	0.03	(0.02)	0.04	(0.06)	-0.01	(0.05)	-0.14	(0.08)	**-0.05**	(0.02)
Malaysia	**-0.15**	(0.04)	**-0.29**	(0.05)	**-0.10**	(0.03)	0.05	(0.04)	**-0.17**	(0.04)	0.03	(0.03)	**-0.14**	(0.02)
Malta	**-0.24**	(0.07)	**-0.26**	(0.06)	**-0.43**	(0.09)	-0.02	(0.06)	**-0.15**	(0.07)	0.02	(0.04)	**-0.41**	(0.11)
Moldova	0.04	(0.05)	**-0.41**	(0.05)	**-0.31**	(0.06)	0.05	(0.05)	**-0.15**	(0.06)	-	-	-	-
Montenegro	**0.21**	(0.06)	**-0.13**	(0.04)	**-0.21**	(0.06)	0.05	(0.04)	-0.03	(0.05)	-0.05	(0.04)	**-0.18**	(0.05)
Morocco	0.10	(0.06)	**-0.32**	(0.06)	-0.48	(0.26)	0.12	(0.08)	**-0.15**	(0.08)	**0.16**	(0.02)	0.08	(0.06)
North Macedonia	0.02	(0.05)	-0.10	(0.05)	**-0.20**	(0.05)	0.03	(0.05)	**-0.34**	(0.05)	-	-	-	-
Panama	**0.09**	(0.04)	**-0.15**	(0.07)	**-0.40**	(0.15)	**-0.22**	(0.06)	-0.04	(0.05)	**0.13**	(0.02)	**-0.17**	(0.07)
Peru	**0.17**	(0.05)	**-0.27**	(0.05)	**-0.19**	(0.06)	-0.05	(0.05)	-0.03	(0.05)	-0.11	(0.08)	**-0.08**	(0.03)
Philippines	**0.16**	(0.08)	**-0.38**	(0.09)	**-0.26**	(0.07)	-0.02	(0.08)	-0.13	(0.09)	**0.24**	(0.03)	**-0.19**	(0.06)
Qatar	**0.14**	(0.04)	**-0.36**	(0.04)	**-0.35**	(0.06)	-0.01	(0.04)	**-0.56**	(0.04)	0.04	(0.03)	**-0.46**	(0.08)
Romania	**0.22**	(0.06)	**-0.20**	(0.04)	**-0.37**	(0.06)	0.07	(0.04)	**-0.27**	(0.05)	-	-	-	-
Saudi Arabia	-0.04	(0.05)	**-0.26**	(0.05)	-0.07	(0.06)	0.02	(0.06)	-0.13	(0.07)	-	-	-	-
Serbia	0.06	(0.05)	**-0.24**	(0.06)	**-0.26**	(0.13)	0.08	(0.05)	**-0.20**	(0.06)	0.04	(0.03)	-0.14	(0.09)
Singapore	-0.03	(0.04)	**-0.21**	(0.03)	-0.01	(0.02)	**-0.17**	(0.04)	**-0.07**	(0.03)	**-0.16**	(0.05)	**-0.19**	(0.04)
Chinese Taipei	**-0.11**	(0.04)	**-0.22**	(0.04)	0.02	(0.02)	**-0.32**	(0.05)	0.02	(0.03)	**-0.19**	(0.04)	**-0.21**	(0.05)
Thailand	0.02	(0.07)	**-0.45**	(0.06)	-0.01	(0.04)	**-0.36**	(0.07)	**0.12**	(0.05)	**0.23**	(0.05)	**-0.09**	(0.04)
Ukraine	**-0.13**	(0.05)	**-0.40**	(0.05)	**-0.49**	(0.08)	0.02	(0.05)	**-0.30**	(0.05)	-	-	-	-
United Arab Emirates	**0.25**	(0.12)	-0.03	(0.04)	**-0.36**	(0.05)	**-0.13**	(0.04)	**-0.43**	(0.04)	0.00	(0.04)	**-0.18**	(0.04)
Uruguay	**0.17**	(0.04)	**-0.19**	(0.05)	**-0.64**	(0.15)	-0.09	(0.06)	-0.08	(0.06)	0.01	(0.03)	**-0.20**	(0.06)
Viet Nam	0.00	(0.07)	-0.12	(0.06)	-0.02	(0.01)	0.07	(0.08)	-0.07	(0.10)	-	-	-	-

Note: Performance decline and Rapid guessing could only be computed for countries/economies with a computer-based assessment.

Annex D. **Differences in disengagement by immigrant background**

Table D.1. Differences between students with and without an immigrant background in PISA 2018 for different measures of disengagement (OECD countries)

Country / Economy	Self-report				Questionnaire-based behaviour						Test-based behaviour			
	Effort (R)		Work mastery (R)		Non-response		Non-dif-ferentiation		Incon-sistency		Performance decline		Rapid guessing	
	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.
Australia	-0.03	(0.03)	**-0.16**	(0.02)	**-0.06**	(0.02)	**-0.07**	(0.03)	**0.05**	(0.02)	-0.02	(0.02)	-0.04	(0.03)
Austria	**0.12**	(0.04)	-0.03	(0.04)	**0.19**	(0.05)	0.04	(0.03)	**0.16**	(0.04)	**0.08**	(0.03)	**0.19**	(0.05)
Belgium	**0.12**	(0.03)	-	-	**0.33**	(0.06)	0.02	(0.04)	**0.21**	(0.04)	**0.21**	(0.03)	**0.21**	(0.04)
Canada	**0.05**	(0.02)	**-0.09**	(0.02)	0.02	(0.03)	**-0.05**	(0.02)	0.02	(0.02)	0.01	(0.02)	0.01	(0.02)
Chile	0.03	(0.08)	-0.09	(0.08)	0.04	(0.09)	0.02	(0.08)	0.01	(0.08)	-0.03	(0.04)	0.02	(0.12)
Colombia	-0.22	(0.15)	-0.03	(0.24)	**0.69**	(0.26)	0.62	(0.35)	0.18	(0.16)	**-0.17**	(0.06)	0.16	(0.29)
Czech Republic	0.08	(0.08)	**0.22**	(0.07)	**0.37**	(0.11)	0.16	(0.09)	**0.54**	(0.17)	0.06	(0.09)	**0.51**	(0.12)
Denmark	**0.24**	(0.04)	**-0.11**	(0.04)	**0.25**	(0.04)	**0.14**	(0.04)	**0.16**	(0.04)	**0.16**	(0.04)	**0.25**	(0.04)
Estonia	0.08	(0.04)	0.01	(0.04)	**0.11**	(0.04)	**0.10**	(0.05)	**0.26**	(0.06)	**0.12**	(0.05)	**0.18**	(0.05)
Finland	0.14	(0.08)	-0.01	(0.06)	**0.33**	(0.10)	0.03	(0.07)	**0.48**	(0.07)	0.07	(0.05)	**0.28**	(0.06)
France	0.09	(0.05)	**-0.15**	(0.04)	**0.45**	(0.08)	-0.07	(0.04)	**0.32**	(0.04)	0.03	(0.04)	**0.32**	(0.06)
Germany	**0.19**	(0.04)	**-0.11**	(0.04)	**0.31**	(0.06)	0.07	(0.04)	**0.22**	(0.04)	0.01	(0.04)	**0.25**	(0.06)
Greece	**-0.12**	(0.04)	**0.12**	(0.04)	**0.21**	(0.06)	0.02	(0.04)	0.07	(0.05)	0.02	(0.03)	**0.26**	(0.08)
Hungary	-0.09	(0.08)	0.07	(0.10)	0.10	(0.08)	-0.06	(0.12)	-0.05	(0.10)	-0.01	(0.13)	0.02	(0.08)
Iceland	0.07	(0.09)	**0.25**	(0.08)	0.17	(0.10)	0.09	(0.09)	0.07	(0.08)	0.03	(0.06)	0.11	(0.11)
Ireland	**0.08**	(0.03)	**-0.08**	(0.03)	0.03	(0.03)	-0.01	(0.03)	0.02	(0.03)	0.01	(0.04)	0.03	(0.02)
Israel	0.07	(0.04)	0.06	(0.04)	**0.26**	(0.07)	**-0.09**	(0.04)	**0.11**	(0.05)	0.01	(0.03)	0.11	(0.06)
Italy	0.05	(0.04)	**0.20**	(0.05)	0.05	(0.06)	**-0.12**	(0.04)	**0.19**	(0.05)	0.05	(0.04)	**0.12**	(0.06)
Japan	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Korea	-0.36	(0.38)	-0.37	(0.58)	0.09	(0.13)	**1.15**	(0.21)	0.01	(0.22)	0.32	(0.17)	**-0.29**	(0.09)
Latvia	0.12	(0.07)	**0.22**	(0.07)	0.09	(0.07)	0.08	(0.07)	0.12	(0.07)	-0.01	(0.07)	0.06	(0.08)
Lithuania	0.10	(0.08)	**0.22**	(0.11)	0.03	(0.06)	-0.05	(0.09)	**0.33**	(0.15)	-0.05	(0.07)	0.20	(0.12)
Luxembourg	0.07	(0.04)	-0.02	(0.03)	**0.12**	(0.03)	**0.07**	(0.03)	**0.14**	(0.03)	**0.07**	(0.03)	**0.15**	(0.02)
Mexico	0.11	(0.09)	**0.57**	(0.20)	**0.39**	(0.17)	**-0.42**	(0.15)	**1.17**	(0.56)	**-0.13**	(0.06)	**-0.08**	(0.03)
Netherlands	**0.22**	(0.06)	**-0.23**	(0.04)	**0.11**	(0.04)	**0.17**	(0.06)	**0.14**	(0.05)	**0.18**	(0.05)	**0.30**	(0.06)
New Zealand	0.01	(0.04)	**-0.18**	(0.03)	-0.02	(0.02)	-0.01	(0.03)	0.01	(0.03)	0.01	(0.03)	-0.07	(0.04)
Norway	**0.13**	(0.04)	-0.02	(0.04)	**0.22**	(0.06)	0.00	(0.05)	**0.21**	(0.04)	0.00	(0.04)	**0.26**	(0.06)
Poland	-0.01	(0.23)	0.33	(0.22)	**0.62**	(0.28)	-0.30	(0.19)	0.60	(0.32)	0.10	(0.20)	0.16	(0.30)
Portugal	0.05	(0.07)	0.07	(0.07)	**0.13**	(0.04)	**-0.24**	(0.05)	**0.22**	(0.06)	0.08	(0.06)	0.13	(0.07)
Slovak Republic	-0.20	(0.18)	0.25	(0.16)	0.12	(0.19)	-0.13	(0.12)	**0.57**	(0.16)	-0.03	(0.08)	**0.36**	(0.17)
Slovenia	0.06	(0.07)	0.03	(0.05)	**0.17**	(0.05)	**0.20**	(0.05)	**0.17**	(0.07)	**0.18**	(0.05)	**0.25**	(0.08)
Spain	**0.04**	(0.02)	**0.08**	(0.02)	0.01	(0.06)	-0.02	(0.02)	0.04	(0.02)	-0.01	(0.01)	0.03	(0.03)
Sweden	**0.15**	(0.05)	**-0.28**	(0.04)	**0.42**	(0.07)	**0.23**	(0.05)	**0.25**	(0.04)	**0.08**	(0.04)	**0.28**	(0.06)
Switzerland	**0.08**	(0.04)	**-0.13**	(0.03)	**0.12**	(0.04)	**0.08**	(0.03)	0.06	(0.04)	**0.14**	(0.03)	**0.26**	(0.04)
Turkey	0.12	(0.10)	0.17	(0.17)	0.16	(0.12)	-0.05	(0.13)	0.10	(0.18)	0.01	(0.15)	0.04	(0.08)
United Kingdom	0.04	(0.04)	**-0.20**	(0.04)	**0.21**	(0.05)	-0.03	(0.04)	**0.15**	(0.03)	0.02	(0.03)	**0.19**	(0.05)
United States	**0.08**	(0.04)	0.03	(0.04)	0.02	(0.02)	-0.04	(0.05)	0.02	(0.04)	0.04	(0.05)	0.05	(0.03)
OECD average	**0.05**	(0.02)	0.02	(0.02)	**0.19**	(0.02)	**0.04**	(0.02)	**0.20**	(0.02)	**0.04**	(0.01)	**0.15**	(0.02)

Note: Question ST182 underlying the computation of Work Mastery has not been administered in Belgium. Question ST019 underlying the computation of IMMIG (Index Immigration Background) was not administered in Japan.

Table D.2. Differences between students with and without an immigrant background in PISA 2018 for different measures of disengagement (Partner countries and economies)

Country / Economy	Self-report				Questionnaire-based behaviour						Test-based behaviour			
	Effort (R)		Work mastery (R)		Non-response		Non-dif-ferentiation		Incon-sistency		Performance decline		Rapid guessing	
	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.
Albania	0.23	(0.13)	0.18	(0.21)	0.18	(0.14)	-0.20	(0.21)	**0.67**	(0.19)	-0.08	(0.11)	**0.33**	(0.16)
Argentina	-0.03	(0.05)	-0.07	(0.05)	0.17	(0.17)	0.10	(0.06)	**-0.17**	(0.06)	-	-	-	-
Baku (Azerbaijan)	0.07	(0.09)	0.02	(0.09)	**0.31**	(0.15)	-0.13	(0.07)	0.12	(0.09)	0.00	(0.06)	0.14	(0.08)
Bosnia and Herzegovina	0.10	(0.15)	0.04	(0.11)	0.22	(0.13)	0.04	(0.09)	0.20	(0.12)	0.01	(0.06)	**0.26**	(0.12)
Brazil	-0.01	(0.19)	0.27	(0.22)	0.80	(0.42)	**-0.48**	(0.21)	**0.59**	(0.22)	-0.18	(0.11)	0.12	(0.16)
Brunei Darussalam	**-0.10**	(0.03)	0.00	(0.03)	**-0.16**	(0.03)	-0.05	(0.03)	0.02	(0.05)	0.03	(0.02)	**-0.07**	(0.02)
B-S-J-Z (China)	-0.17	(0.21)	-0.30	(0.49)	0.07	(0.05)	-0.34	(0.43)	-0.07	(0.17)	**0.88**	(0.23)	-0.04	(0.03)
Bulgaria	0.18	(0.15)	**0.91**	(0.13)	**0.59**	(0.27)	-0.23	(0.20)	**1.37**	(0.30)	-0.05	(0.09)	**0.86**	(0.29)
Costa Rica	-0.03	(0.04)	0.00	(0.05)	**0.10**	(0.05)	0.00	(0.05)	0.03	(0.05)	**-0.03**	(0.01)	**0.18**	(0.08)
Croatia	0.05	(0.06)	0.00	(0.04)	0.07	(0.05)	0.00	(0.05)	0.06	(0.05)	0.01	(0.04)	0.08	(0.04)
Dominican Republic	0.03	(0.10)	0.25	(0.18)	**0.57**	(0.25)	0.03	(0.15)	0.09	(0.20)	-0.04	(0.03)	-0.07	(0.17)
Georgia	0.27	(0.21)	0.27	(0.19)	**0.87**	(0.31)	0.26	(0.22)	**0.79**	(0.22)	-0.17	(0.19)	**0.64**	(0.27)
Hong Kong (China)	**-0.08**	(0.03)	-0.02	(0.03)	0.01	(0.01)	0.02	(0.03)	-0.04	(0.03)	0.05	(0.04)	-0.02	(0.03)
Indonesia	**-0.30**	(0.13)	0.36	(0.29)	0.07	(0.09)	-0.78	(0.61)	**1.90**	(0.52)	-0.05	(0.03)	0.08	(0.15)
Jordan	0.04	(0.05)	-0.11	(0.06)	0.00	(0.05)	0.00	(0.05)	-0.08	(0.07)	-	-	-	-
Kazakhstan	-0.04	(0.04)	0.00	(0.04)	0.03	(0.04)	-0.04	(0.05)	0.04	(0.04)	-0.03	(0.02)	0.02	(0.03)
Kosovo	0.22	(0.19)	-0.02	(0.16)	0.20	(0.17)	0.30	(0.20)	0.00	(0.20)	0.02	(0.06)	0.06	(0.10)
Lebanon	0.07	(0.10)	**0.23**	(0.08)	0.15	(0.14)	-0.16	(0.10)	0.21	(0.12)	-	-	-	-
Macao (China)	0.00	(0.03)	**-0.07**	(0.03)	0.00	(0.01)	-0.03	(0.03)	-0.03	(0.03)	-0.10	(0.05)	**-0.04**	(0.01)
Malaysia	0.03	(0.12)	0.18	(0.10)	0.07	(0.10)	**-0.27**	(0.11)	**0.39**	(0.17)	-0.02	(0.11)	0.05	(0.10)
Malta	**0.18**	(0.09)	**0.23**	(0.08)	**0.22**	(0.10)	-0.13	(0.07)	**0.21**	(0.07)	0.00	(0.05)	0.14	(0.12)
Moldova	**0.51**	(0.21)	0.08	(0.12)	0.04	(0.11)	-0.13	(0.13)	0.27	(0.17)	-	-	-	-
Montenegro	0.09	(0.07)	0.05	(0.06)	0.08	(0.08)	-0.01	(0.06)	**0.22**	(0.07)	0.05	(0.03)	-0.01	(0.07)
Morocco	0.29	(0.20)	**0.75**	(0.17)	0.00	(0.30)	-0.28	(0.17)	**1.16**	(0.26)	**-0.12**	(0.04)	**0.62**	(0.25)
North Macedonia	0.22	(0.21)	0.17	(0.18)	**0.50**	(0.23)	-0.26	(0.16)	**0.42**	(0.18)	-	-	-	-
Panama	0.09	(0.06)	0.11	(0.06)	**-0.28**	(0.11)	-0.20	(0.11)	0.16	(0.13)	0.03	(0.02)	-0.08	(0.10)
Peru	**-0.29**	(0.10)	0.28	(0.39)	0.53	(0.30)	-0.23	(0.30)	**1.12**	(0.50)	0.17	(0.10)	0.30	(0.22)
Philippines	**0.54**	(0.24)	**0.97**	(0.17)	**0.62**	(0.24)	**-1.16**	(0.17)	**1.88**	(0.24)	**-0.12**	(0.03)	**0.44**	(0.18)
Qatar	**0.07**	(0.02)	**-0.19**	(0.02)	**-0.45**	(0.03)	**-0.15**	(0.02)	**-0.41**	(0.03)	0.01	(0.02)	**-0.82**	(0.04)
Romania	-0.04	(0.25)	0.09	(0.24)	0.14	(0.21)	**-0.40**	(0.16)	**0.56**	(0.25)	-	-	-	-
Saudi Arabia	0.03	(0.05)	**-0.21**	(0.04)	0.02	(0.08)	-0.05	(0.06)	**-0.20**	(0.05)	-	-	-	-
Serbia	0.05	(0.07)	-0.03	(0.06)	-0.09	(0.09)	-0.02	(0.06)	0.06	(0.06)	0.00	(0.03)	0.04	(0.10)
Singapore	0.01	(0.03)	**-0.09**	(0.03)	0.01	(0.01)	**-0.07**	(0.03)	-0.02	(0.02)	-0.04	(0.03)	-0.04	(0.03)
Chinese Taipei	0.44	(0.26)	0.20	(0.21)	**0.98**	(0.43)	-0.04	(0.13)	0.42	(0.35)	**0.41**	(0.16)	0.57	(0.35)
Thailand	**0.32**	(0.10)	0.35	(0.20)	0.16	(0.14)	**-0.51**	(0.16)	**0.58**	(0.20)	0.06	(0.07)	0.26	(0.14)
Ukraine	0.23	(0.13)	0.11	(0.11)	0.10	(0.11)	-0.05	(0.10)	**0.48**	(0.10)	-	-	-	-
United Arab Emirates	**0.13**	(0.05)	**-0.07**	(0.02)	**-0.30**	(0.03)	**-0.23**	(0.03)	**-0.30**	(0.02)	-0.01	(0.03)	**-0.42**	(0.03)
Uruguay	-0.21	(0.14)	-0.26	(0.17)	0.51	(0.34)	**-0.37**	(0.14)	**0.49**	(0.21)	-0.33	(0.18)	0.07	(0.19)
Viet Nam	**-0.33**	(0.16)	**1.36**	(0.27)	**-0.03**	(0.01)	**-0.73**	(0.06)	1.23	(0.63)	-	-	-	-

Note: Performance decline and Rapid guessing could only be computed for countries/economies with a computer-based assessment.

Annex E. Relationship between disengagement and performance

Table E.1. Relationship between different measures of disengagement and reading performance in PISA 2018 (OECD countries)

Country / Economy	Self-report				Questionnaire-based behaviour						Test-based behaviour	
	Effort (R)		Work mastery (R)		Non-response		Non-dif-ferentiation		Inconsistency		Rapid guessing	
	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.
Australia	**-22.30**	(1.22)	**-12.94**	(1.21)	**-28.19**	(1.79)	**-24.35**	(1.15)	**-4.53**	(1.52)	**-43.90**	(1.44)
Austria	**-13.43**	(1.36)	**-9.61**	(1.32)	**-28.89**	(1.82)	**-21.92**	(1.54)	**-13.14**	(1.53)	**-41.73**	(3.89)
Belgium	**-15.26**	(1.56)	-	-	**-22.68**	(1.63)	**-12.47**	(1.45)	**-15.43**	(1.36)	**-31.00**	(2.67)
Canada	**-16.32**	(1.05)	**-9.80**	(1.12)	**-24.21**	(1.54)	**-20.14**	(1.07)	**-5.31**	(1.19)	**-38.43**	(1.99)
Chile	**-3.22**	(1.50)	**-9.14**	(1.23)	**-10.72**	(1.88)	**-14.80**	(1.33)	**-7.48**	(1.26)	**-12.98**	(1.69)
Colombia	-2.98	(1.73)	**-10.84**	(1.20)	**-9.67**	(1.12)	**-10.62**	(1.13)	**-7.71**	(1.24)	**-4.76**	(1.27)
Czech Republic	**-11.00**	(1.46)	**-9.81**	(1.53)	**-18.41**	(2.47)	**-13.65**	(1.31)	**-15.19**	(1.32)	**-33.26**	(1.76)
Denmark	**-23.36**	(1.68)	**-13.87**	(1.65)	**-26.08**	(2.46)	**-14.93**	(1.62)	**-13.72**	(2.09)	**-36.15**	(3.10)
Estonia	**-12.35**	(1.79)	**-10.09**	(1.71)	**-30.37**	(2.33)	**-14.26**	(1.54)	**-10.89**	(1.63)	**-31.82**	(2.41)
Finland	**-33.46**	(1.58)	**-21.11**	(1.30)	**-26.28**	(2.13)	**-22.56**	(1.31)	**-14.44**	(1.49)	**-39.06**	(2.39)
France	**-8.21**	(1.20)	**-7.54**	(1.23)	**-19.85**	(1.32)	**-15.91**	(1.48)	**-12.91**	(1.35)	**-26.06**	(1.84)
Germany	**-17.34**	(1.62)	**-7.81**	(1.72)	**-27.00**	(1.54)	**-20.23**	(1.71)	**-12.80**	(1.72)	**-30.46**	(2.42)
Greece	**-8.17**	(1.29)	**-17.12**	(1.35)	**-18.75**	(2.00)	**-11.44**	(1.22)	**-20.05**	(1.19)	**-24.82**	(1.38)
Hungary	**-10.28**	(1.59)	**-5.88**	(1.31)	**-19.87**	(2.01)	**-20.94**	(1.53)	**-4.26**	(1.55)	**-31.57**	(2.30)
Iceland	**-26.95**	(1.61)	**-16.65**	(1.61)	**-21.63**	(1.78)	**-25.80**	(1.84)	**-7.03**	(2.32)	**-39.31**	(2.07)
Ireland	**-8.69**	(1.63)	**-6.42**	(1.38)	**-30.61**	(2.37)	**-18.76**	(1.45)	**-8.80**	(1.58)	**-37.64**	(2.87)
Israel	**-12.73**	(1.68)	**-13.95**	(1.75)	**-26.53**	(1.35)	**-16.72**	(1.47)	**-16.57**	(1.48)	**-39.91**	(1.54)
Italy	**-20.80**	(1.78)	**-13.14**	(1.35)	**-16.49**	(1.07)	**-12.39**	(1.98)	**-18.50**	(1.67)	**-32.79**	(1.57)
Japan	1.90	(1.52)	-0.85	(1.44)	**-42.95**	(4.45)	**-25.51**	(1.48)	**-12.83**	(1.71)	**-42.30**	(2.92)
Korea	**-21.76**	(1.57)	**-13.90**	(1.81)	**-53.90**	(7.31)	**-34.54**	(1.65)	**10.79**	(2.52)	**-39.12**	(1.09)
Latvia	**-23.53**	(1.59)	**-12.00**	(1.23)	**-24.12**	(2.99)	**-13.96**	(1.46)	**-13.56**	(1.31)	**-30.81**	(2.75)
Lithuania	**-19.25**	(1.44)	**-2.75**	(1.27)	**-23.60**	(1.51)	**-15.36**	(1.11)	**-7.65**	(1.27)	**-35.37**	(1.79)
Luxembourg	**-15.08**	(1.26)	**-10.00**	(1.46)	**-29.67**	(1.92)	**-15.06**	(1.29)	**-20.12**	(1.22)	**-41.72**	(2.20)
Mexico	**-7.09**	(2.14)	**-12.17**	(1.37)	**-9.31**	(1.88)	**-7.38**	(1.26)	**-8.21**	(1.42)	**-13.38**	(3.01)
Netherlands	**-38.41**	(2.01)	1.64	(2.17)	**-24.96**	(5.40)	**-22.22**	(2.31)	**-17.00**	(2.50)	**-54.02**	(3.01)
New Zealand	**-19.02**	(1.56)	**-13.65**	(1.86)	**-35.21**	(3.27)	**-24.61**	(1.67)	**-3.72**	(1.84)	**-36.96**	(2.08)
Norway	**-26.06**	(1.46)	**-11.09**	(1.39)	**-30.51**	(1.87)	**-21.21**	(1.37)	**-22.61**	(1.27)	**-34.74**	(1.77)
Poland	**-7.81**	(1.48)	**-14.48**	(1.47)	**-21.49**	(2.33)	**-16.51**	(1.37)	**-10.90**	(1.75)	**-36.39**	(2.21)
Portugal	**-9.53**	(1.55)	**-8.12**	(1.44)	**-16.51**	(1.79)	**-15.10**	(1.36)	-0.88	(1.55)	**-15.75**	(2.46)
Slovak Republic	**-9.12**	(1.12)	**-12.09**	(1.33)	**-11.88**	(1.39)	**-10.47**	(1.23)	**-13.27**	(1.29)	**-24.61**	(1.56)
Slovenia	**-15.08**	(1.36)	**-16.81**	(1.39)	**-25.56**	(2.37)	**-15.19**	(1.61)	**-14.84**	(1.13)	**-41.28**	(1.72)
Spain	**-12.63**	(0.97)	**-7.25**	(0.82)	**-10.62**	(1.14)	**-9.21**	(0.89)	**-14.82**	(0.89)	**-26.52**	(1.59)
Sweden	**-18.05**	(1.64)	**-13.22**	(1.60)	**-27.12**	(2.00)	**-21.72**	(1.52)	**-18.29**	(1.61)	**-20.44**	(1.99)
Switzerland	**-17.36**	(1.87)	**-8.17**	(1.97)	**-20.16**	(1.84)	**-16.65**	(1.70)	**-15.71**	(1.89)	**-34.70**	(2.34)
Turkey	1.68	(1.91)	**-3.18**	(1.00)	**-29.55**	(4.32)	**-14.36**	(1.21)	**-8.03**	(1.28)	**-36.29**	(4.08)
United Kingdom	**-15.64**	(1.80)	**-10.74**	(1.55)	**-29.97**	(1.59)	**-18.89**	(1.72)	**-13.71**	(1.46)	**-29.24**	(1.25)
United States	**-15.40**	(2.85)	**-8.26**	(1.67)	**-31.94**	(3.98)	**-20.64**	(1.98)	**-17.19**	(1.81)	**-50.85**	(2.67)
OECD average	**-14.98**	(0.27)	**-10.36**	(0.25)	**-24.47**	(0.44)	**-17.58**	(0.25)	**-11.66**	(0.26)	**-32.98**	(0.38)

Note: Question ST182 underlying the computation of Work Mastery has not been administered in Belgium.

Table E.2. Relationship between different measures of disengagement and reading performance in PISA 2018 (Partner countries and economies)

Country / Economy	Self-report				Questionnaire-based behaviour						Test-based behaviour	
	Effort (R)		Work mastery (R)		Non-response		Non-dif-ferentiation		Inconsistency		Rapid guessing	
	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.
Albania	**-12.26**	(1.90)	**-13.39**	(1.29)	**-22.92**	(1.53)	**-6.44**	(1.07)	**-12.07**	(1.06)	**-29.40**	(1.49)
Argentina	**-12.21**	(1.68)	**-11.29**	(1.39)	**-9.01**	(0.80)	**-7.99**	(1.20)	**-12.65**	(1.06)	-	-
Baku (Azerbaijan)	-2.06	(1.13)	**-5.91**	(0.95)	**-10.71**	(0.76)	**-6.81**	(0.82)	**-5.79**	(0.84)	**-21.02**	(0.97)
Bosnia and Herzegovina	-0.66	(1.25)	**-13.09**	(1.18)	**-18.07**	(0.97)	**-9.32**	(0.91)	**-10.58**	(0.73)	**-22.65**	(1.02)
Brazil	**-11.53**	(1.46)	**-10.73**	(1.24)	**-8.22**	(0.74)	**-8.34**	(0.97)	**-7.75**	(0.85)	**-12.69**	(0.99)
Brunei Darussalam	**-16.57**	(1.46)	**-20.83**	(1.28)	**-23.23**	(1.44)	**-8.54**	(1.12)	**-11.15**	(1.18)	**-26.43**	(1.99)
B-S-J-Z (China)	0.92	(1.76)	**-6.29**	(1.30)	**-96.00**	(17.46)	**-21.00**	(1.28)	**-7.01**	(1.70)	**-37.54**	(5.53)
Bulgaria	**-6.04**	(1.44)	**-16.05**	(1.71)	**-11.57**	(1.05)	**-7.87**	(1.22)	**-11.93**	(0.97)	**-26.65**	(1.07)
Costa Rica	-1.85	(1.29)	**-4.52**	(0.93)	**-21.62**	(1.59)	**-14.36**	(1.24)	**-7.93**	(1.08)	**-16.72**	(2.81)
Croatia	**-10.13**	(1.40)	**-6.00**	(1.59)	**-24.30**	(1.52)	**-15.30**	(1.27)	**-10.86**	(1.43)	**-36.81**	(1.52)
Dominican Republic	-0.29	(1.66)	**-11.24**	(1.14)	**-5.34**	(0.71)	**-6.59**	(0.78)	**-4.23**	(0.85)	0.43	(0.73)
Georgia	**2.76**	(1.19)	**-16.76**	(1.14)	**-18.01**	(0.94)	**-6.36**	(1.22)	**-10.75**	(1.03)	**-20.23**	(0.86)
Hong Kong (China)	**-12.84**	(1.51)	**-4.59**	(1.68)	**-40.16**	(6.16)	**-22.89**	(1.08)	0.48	(1.63)	**-45.42**	(1.63)
Indonesia	**-6.52**	(1.98)	**-17.13**	(1.26)	**-10.84**	(3.27)	**-7.61**	(1.16)	**-7.83**	(1.36)	**-19.31**	(3.59)
Jordan	**-5.08**	(1.62)	**-20.79**	(1.07)	**-20.44**	(1.37)	**-3.00**	(1.09)	**-11.59**	(0.93)	-	-
Kazakhstan	**3.35**	(1.00)	**-3.05**	(0.71)	**-17.11**	(1.04)	**-10.72**	(0.71)	**-3.95**	(0.67)	**-23.26**	(1.23)
Kosovo	**-2.99**	(1.08)	**-21.32**	(1.16)	**-15.18**	(0.82)	**-8.41**	(0.92)	**-10.05**	(1.04)	**-19.51**	(1.48)
Lebanon	**-8.21**	(2.08)	**-24.30**	(2.14)	**-7.16**	(1.51)	**8.96**	(2.10)	**-18.46**	(1.60)	-	-
Macao (China)	**-17.54**	(1.88)	**-5.22**	(1.52)	-22.36	(7.04)	**-14.48**	(1.56)	-2.01	(1.80)	**-40.61**	(4.35)
Malaysia	**-16.89**	(1.41)	**-18.67**	(1.38)	**-13.16**	(3.73)	**-3.71**	(1.11)	**-11.70**	(1.13)	**-21.85**	(2.68)
Malta	**-20.81**	(1.66)	**-20.74**	(1.80)	**-26.17**	(1.24)	**-18.80**	(1.55)	**-15.92**	(1.77)	**-30.81**	(1.14)
Moldova	**-10.30**	(2.34)	**-15.09**	(1.72)	**-23.17**	(1.49)	**-4.81**	(1.83)	**-12.23**	(1.89)	-	-
Montenegro	**-4.31**	(0.99)	**-8.22**	(1.02)	**-20.06**	(0.71)	**-11.43**	(1.11)	**-10.41**	(0.77)	**-28.79**	(0.89)
Morocco	-2.07	(1.08)	**-12.58**	(0.93)	**-4.10**	(0.59)	-0.77	(0.69)	**-4.70**	(0.81)	**-12.50**	(1.56)
North Macedonia	**-14.13**	(1.51)	**-6.85**	(1.34)	**-25.32**	(1.42)	**-9.71**	(1.06)	**-5.14**	(1.11)	-	-
Panama	-2.39	(1.80)	**-10.20**	(1.56)	**-8.71**	(0.81)	**-4.01**	(1.02)	**-9.26**	(0.93)	**-10.30**	(1.21)
Peru	-0.50	(1.57)	**-14.35**	(1.06)	**-16.88**	(2.48)	**-5.29**	(1.00)	**-8.05**	(0.92)	**-22.60**	(2.54)
Philippines	**-5.25**	(1.43)	**-20.71**	(1.25)	**-8.75**	(1.35)	**-2.42**	(1.00)	**-7.92**	(0.83)	**-12.10**	(1.22)
Qatar	**-5.69**	(1.02)	**-14.98**	(0.82)	**-14.78**	(0.50)	**-8.56**	(0.71)	**-14.63**	(0.68)	**-21.65**	(0.42)
Romania	-2.98	(1.96)	**-7.67**	(1.95)	**-17.68**	(1.59)	**-6.25**	(1.72)	**-16.36**	(1.61)	-	-
Saudi Arabia	**-6.89**	(1.57)	**-13.58**	(1.19)	**-9.90**	(1.24)	**-6.38**	(1.05)	**-10.57**	(1.13)	-	-
Serbia	**-8.17**	(1.51)	**-12.68**	(1.63)	**-17.15**	(0.88)	**-11.47**	(1.39)	**-11.18**	(1.21)	**-29.03**	(1.12)
Singapore	**-20.06**	(1.37)	**-4.26**	(1.65)	**-42.01**	(4.04)	**-23.27**	(1.38)	-1.36	(1.72)	**-51.59**	(1.88)
Chinese Taipei	**-24.31**	(1.25)	**-5.06**	(1.55)	**-36.98**	(3.56)	**-32.69**	(1.20)	**3.76**	(1.64)	**-39.48**	(1.35)
Thailand	**-15.23**	(1.60)	**-21.32**	(1.24)	**-20.35**	(2.07)	**-11.29**	(1.08)	**-3.00**	(1.35)	**-25.59**	(1.77)
Ukraine	**-8.77**	(2.11)	**-4.20**	(1.90)	**-20.77**	(1.63)	**-11.40**	(1.73)	**-14.10**	(1.21)	-	-
United Arab Emirates	**-8.85**	(2.78)	**-14.06**	(0.84)	**-16.00**	(0.65)	**-9.82**	(0.74)	**-15.05**	(0.66)	**-31.72**	(0.91)
Uruguay	**-6.46**	(1.70)	**-6.68**	(1.39)	**-15.13**	(0.86)	**-9.31**	(1.08)	**-10.31**	(1.10)	**-12.75**	(1.98)

Note: Rapid guessing could only be computed for countries/economies with a computer-based assessment (CBA).

Annex F. Stability across cycles for Self-reported effort

Table F.1. Country-level means and difference across cycles for Self-reported effort (OECD countries)

Country / Economy	Effort (R)					
	2018		2012		Difference	
	Mean	S.E.	Mean	S.E.	Est.	S.E.
Australia	**3.57**	(0.03)	**3.44**	(0.02)	**0.13**	(0.03)
Austria	**3.85**	(0.03)	**3.61**	(0.04)	**0.24**	(0.05)
Belgium	**3.72**	(0.02)	**3.76**	(0.03)	-0.04	(0.04)
Canada	**3.53**	(0.02)	**4.27**	(0.02)	**-0.75**	(0.03)
Chile	**3.00**	(0.03)	**2.83**	(0.04)	**0.18**	(0.05)
Colombia	**2.53**	(0.04)	**2.52**	(0.03)	0.02	(0.05)
Czech Republic	**3.73**	(0.03)	**3.52**	(0.03)	**0.22**	(0.05)
Denmark	**3.50**	(0.03)	**3.37**	(0.04)	**0.14**	(0.05)
Estonia	**3.28**	(0.03)	**3.69**	(0.04)	**-0.41**	(0.05)
Finland	**3.02**	(0.03)	**3.00**	(0.03)	0.02	(0.05)
France	**3.84**	(0.04)	**4.32**	(0.04)	**-0.48**	(0.06)
Germany	**3.83**	(0.04)	**3.96**	(0.04)	**-0.13**	(0.05)
Hungary	**3.30**	(0.04)	**3.37**	(0.04)	-0.07	(0.06)
Ireland	**3.02**	(0.03)	**3.09**	(0.03)	-0.07	(0.04)
Israel	**3.13**	(0.04)	**3.38**	(0.04)	**-0.26**	(0.06)
Italy	**3.05**	(0.03)	**3.09**	(0.04)	-0.04	(0.06)
Japan	**3.86**	(0.04)	**4.71**	(0.05)	**-0.85**	(0.07)
Korea	**2.74**	(0.03)	**3.55**	(0.05)	**-0.82**	(0.06)
Netherlands	**3.55**	(0.04)	**3.78**	(0.04)	**-0.23**	(0.06)
Norway	**3.62**	(0.04)	**3.84**	(0.06)	**-0.21**	(0.07)
Poland	**3.56**	(0.04)	**3.50**	(0.05)	0.06	(0.06)
Portugal	**3.50**	(0.03)	**3.34**	(0.04)	**0.16**	(0.05)
Slovak Republic	**3.68**	(0.03)	**3.84**	(0.05)	**-0.16**	(0.06)
Slovenia	**3.44**	(0.03)	**3.66**	(0.03)	**-0.21**	(0.04)
Spain	**3.45**	(0.02)	**3.66**	(0.04)	**-0.21**	(0.04)
Sweden	**3.60**	(0.04)	**3.97**	(0.05)	**-0.36**	(0.06)
Türkiye	**2.09**	(0.04)	**2.21**	(0.04)	**-0.12**	(0.05)
United Kingdom	**3.54**	(0.03)	**3.50**	(0.05)	0.03	(0.06)
United States	**2.75**	(0.03)	**2.54**	(0.03)	**0.21**	(0.04)

Notes: Means refer to the unstandardised measures, with higher values indicating higher levels of disengagement. Only countries/economies are displayed in which self-reported effort was administered in both cycles.

Table F.2. Country-level means and difference across cycles for Self-reported effort (Partner countries and economies)

Country / Economy	Effort (R)					
	2018		2012		Difference	
	Mean	S.E.	Mean	S.E.	Est.	S.E.
Brazil	**2.82**	(0.03)	**2.71**	(0.05)	**0.12**	(0.06)
Bulgaria	**3.28**	(0.04)	**3.02**	(0.04)	**0.26**	(0.06)
Croatia	**3.39**	(0.04)	**3.55**	(0.04)	**-0.16**	(0.05)
Hong Kong (China)	**3.60**	(0.03)	**3.65**	(0.04)	-0.05	(0.05)
Macao (China)	**2.89**	(0.02)	**3.04**	(0.02)	**-0.15**	(0.03)
Malaysia	**2.42**	(0.03)	**3.21**	(0.04)	**-0.79**	(0.05)
Montenegro	**3.27**	(0.03)	**3.50**	(0.04)	**-0.23**	(0.05)
Serbia	**3.58**	(0.04)	**3.67**	(0.05)	-0.09	(0.07)
Singapore	**3.47**	(0.03)	**3.45**	(0.02)	0.02	(0.04)
Chinese Taipei	**2.71**	(0.04)	**2.66**	(0.04)	0.04	(0.05)
United Arab Emirates	**2.49**	(0.09)	**3.18**	(0.03)	**-0.69**	(0.09)
Uruguay	**3.13**	(0.04)	**3.21**	(0.04)	-0.08	(0.05)

Notes: Means refer to the unstandardised measures, with higher values indicating higher levels of disengagement. Only countries/economies are displayed in which self-reported effort was administered in both cycles.

Annex G. Stability across cycles for Non-response

Table G.1. Country-level means and difference across cycles for Non-response (OECD countries)

Country / Economy	2018		2015		2012		2018-2015		2018-2012		2015-2012	
	Mean	S.E.	Mean	S.E.	Mean	S.E.	Diff.	S.E.	Diff.	S.E.	Diff.	S.E.
Australia	**0.04**	(0.002)	**0.06**	(0.002)	**0.03**	(0.001)	**-0.02**	(0.003)	0.00	(0.002)	**0.02**	(0.002)
Austria	**0.04**	(0.003)	**0.05**	(0.002)	**0.03**	(0.002)	**-0.01**	(0.003)	**0.01**	(0.003)	**0.02**	(0.002)
Belgium	**0.04**	(0.002)	**0.06**	(0.003)	**0.04**	(0.001)	**-0.02**	(0.004)	0.00	(0.003)	**0.02**	(0.003)
Canada	**0.04**	(0.002)	**0.04**	(0.002)	**0.03**	(0.001)	0.00	(0.003)	**0.00**	(0.002)	**0.01**	(0.002)
Chile	**0.06**	(0.004)	**0.06**	(0.004)	**0.03**	(0.001)	0.00	(0.006)	**0.03**	(0.004)	**0.04**	(0.004)
Colombia	**0.09**	(0.008)	**0.04**	(0.002)	**0.10**	(0.006)	**0.06**	(0.009)	-0.01	(0.010)	**-0.07**	(0.006)
Costa Rica	**0.04**	(0.003)	**0.07**	(0.004)	-	-	**-0.03**	(0.005)	-	-	-	-
Czech Republic	**0.04**	(0.003)	**0.04**	(0.002)	**0.03**	(0.001)	0.00	(0.004)	**0.01**	(0.003)	**0.01**	(0.003)
Denmark	**0.03**	(0.002)	**0.03**	(0.002)	**0.03**	(0.001)	0.00	(0.003)	**0.00**	(0.002)	0.00	(0.003)
Estonia	**0.02**	(0.001)	**0.02**	(0.001)	**0.03**	(0.001)	0.00	(0.002)	**0.00**	(0.002)	**-0.01**	(0.001)
Finland	**0.03**	(0.002)	**0.03**	(0.002)	**0.03**	(0.001)	0.00	(0.003)	**0.01**	(0.002)	**0.01**	(0.002)
France	**0.06**	(0.004)	**0.06**	(0.004)	**0.05**	(0.002)	0.00	(0.005)	**0.02**	(0.004)	**0.01**	(0.004)
Germany	**0.07**	(0.004)	**0.08**	(0.003)	**0.05**	(0.002)	**-0.01**	(0.005)	**0.02**	(0.004)	**0.03**	(0.004)
Greece	**0.05**	(0.004)	**0.04**	(0.004)	-	-	0.01	(0.006)	-	-	-	-
Hungary	**0.03**	(0.002)	**0.04**	(0.002)	**0.03**	(0.001)	**-0.01**	(0.003)	**0.01**	(0.003)	**0.01**	(0.003)
Iceland	**0.05**	(0.002)	**0.05**	(0.002)	-	-	0.00	(0.003)	-	-	-	-
Ireland	**0.03**	(0.002)	**0.03**	(0.001)	**0.03**	(0.001)	0.00	(0.002)	0.00	(0.002)	0.00	(0.002)
Israel	**0.09**	(0.006)	**0.05**	(0.004)	**0.05**	(0.004)	**0.04**	(0.007)	**0.04**	(0.007)	0.00	(0.006)
Italy	**0.07**	(0.004)	**0.03**	(0.002)	**0.03**	(0.001)	**0.03**	(0.004)	**0.04**	(0.004)	0.00	(0.003)
Japan	**0.02**	(0.001)	**0.02**	(0.003)	**0.02**	(0.001)	0.00	(0.004)	**0.00**	(0.001)	**0.01**	(0.003)
Korea	**0.01**	(0.001)	**0.02**	(0.001)	**0.02**	(0.001)	**-0.01**	(0.001)	**-0.01**	(0.001)	**-0.01**	(0.001)
Latvia	**0.03**	(0.002)	**0.03**	(0.001)	-	-	0.00	(0.002)	-	-	-	-
Lithuania	**0.03**	(0.002)	**0.04**	(0.003)	-	-	0.00	(0.004)	-	-	-	-
Luxembourg	**0.04**	(0.001)	**0.06**	(0.002)	-	-	**-0.02**	(0.002)	-	-	-	-
Mexico	**0.05**	(0.005)	**0.04**	(0.003)	-	-	**0.01**	(0.005)	-	-	-	-
Netherlands	**0.02**	(0.002)	**0.03**	(0.002)	**0.04**	(0.001)	**-0.01**	(0.003)	**-0.01**	(0.002)	**-0.01**	(0.002)
New Zealand	**0.03**	(0.001)	**0.05**	(0.002)	-	-	**-0.02**	(0.003)	-	-	-	-
Norway	**0.04**	(0.003)	**0.05**	(0.003)	**0.04**	(0.002)	0.00	(0.004)	0.01	(0.003)	**0.01**	(0.003)
Poland	**0.02**	(0.002)	**0.02**	(0.001)	**0.01**	(0.001)	**0.01**	(0.002)	**0.01**	(0.002)	**0.00**	(0.001)
Portugal	**0.03**	(0.002)	**0.03**	(0.001)	**0.04**	(0.002)	0.00	(0.003)	-0.01	(0.003)	**-0.01**	(0.002)
Slovak Republic	**0.06**	(0.003)	**0.05**	(0.003)	**0.03**	(0.002)	0.01	(0.004)	**0.02**	(0.004)	**0.02**	(0.004)
Slovenia	**0.02**	(0.001)	**0.03**	(0.001)	**0.02**	(0.001)	**0.00**	(0.001)	0.00	(0.001)	**0.00**	(0.001)
Spain	**0.06**	(0.004)	**0.03**	(0.001)	**0.04**	(0.001)	**0.03**	(0.004)	**0.02**	(0.004)	**-0.01**	(0.002)
Sweden	**0.04**	(0.003)	**0.06**	(0.004)	**0.04**	(0.001)	**-0.02**	(0.005)	0.00	(0.003)	**0.02**	(0.005)
Switzerland	**0.06**	(0.003)	**0.08**	(0.004)	-	-	**-0.03**	(0.005)	-	-	-	-
Türkiye	**0.02**	(0.001)	**0.03**	(0.004)	**0.04**	(0.002)	**-0.01**	(0.004)	**-0.02**	(0.002)	0.00	(0.004)
United Kingdom	**0.05**	(0.002)	**0.05**	(0.003)	**0.04**	(0.001)	0.00	(0.004)	**0.01**	(0.003)	**0.01**	(0.003)
United States	**0.02**	(0.002)	**0.04**	(0.003)	**0.04**	(0.002)	**-0.01**	(0.003)	**-0.02**	(0.003)	0.00	(0.003)

Note: Means refer to the unstandardised measures, with higher values indicating higher levels of disengagement.

Table G.2. Country-level means and difference across cycles for Non-response (Partner countries and economies)

Country / Economy	2018		2015		2012		2018-2015		2018-2012		2015-2012	
	Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.
Brazil	**0.15**	(0.006)	**0.19**	(0.005)	**0.07**	(0.003)	**-0.04**	(0.008)	**0.08**	(0.007)	**0.12**	(0.006)
Bulgaria	**0.14**	(0.008)	**0.10**	(0.007)	**0.07**	(0.004)	**0.04**	(0.010)	**0.07**	(0.009)	**0.03**	(0.008)
Croatia	**0.04**	(0.003)	**0.03**	(0.002)	**0.03**	(0.001)	**0.01**	(0.003)	**0.02**	(0.003)	0.00	(0.002)
Dominican Republic	**0.23**	(0.011)	**0.11**	(0.006)	-	-	**0.12**	(0.013)	-	-	-	-
Georgia	**0.11**	(0.004)	-	-	-	-	-	-	-	-	-	-
Hong Kong (China)	**0.01**	(0.001)	**0.02**	(0.001)	**0.01**	(0.001)	**-0.01**	(0.001)	**0.00**	(0.001)	**0.00**	(0.001)
Macao (China)	**0.01**	(0.000)	**0.01**	(0.000)	**0.02**	(0.000)	**0.00**	(0.000)	**-0.01**	(0.001)	**-0.01**	(0.001)
Malaysia	**0.01**	(0.002)	-	-	**0.04**	(0.003)	-	-	**-0.03**	(0.003)	-	-
Montenegro	**0.09**	(0.002)	**0.09**	(0.002)	**0.05**	(0.001)	0.00	(0.003)	**0.04**	(0.002)	**0.04**	(0.002)
Peru	**0.04**	(0.002)	**0.04**	(0.002)	-	-	0.00	(0.003)	-	-	-	-
Qatar	**0.09**	(0.001)	**0.12**	(0.002)	-	-	**-0.03**	(0.002)	-	-	-	-
Serbia	**0.11**	(0.007)	-	-	**0.04**	(0.002)	-	-	**0.07**	(0.008)	-	-
Singapore	**0.01**	(0.001)	**0.02**	(0.000)	**0.02**	(0.000)	**0.00**	(0.001)	**-0.01**	(0.001)	**0.00**	(0.001)
Chinese Taipei	**0.01**	(0.001)	**0.01**	(0.000)	**0.02**	(0.000)	0.00	(0.001)	**-0.01**	(0.001)	**-0.01**	(0.000)
Thailand	**0.01**	(0.002)	**0.02**	(0.002)	-	-	**-0.01**	(0.002)	-	-	-	-
United Arab Emirates	**0.05**	(0.002)	**0.05**	(0.003)	**0.05**	(0.002)	0.00	(0.003)	0.00	(0.003)	0.00	(0.003)
Uruguay	**0.14**	(0.006)	**0.11**	(0.004)	**0.08**	(0.004)	**0.03**	(0.007)	**0.06**	(0.007)	**0.03**	(0.006)

Notes: Means refer to the unstandardised measures, with higher values indicating higher levels of disengagement. Only countries/economies are displayed for which the information was available in at least two cycles.

Annex H. Stability across cycles for Rapid Guessing

Table H.1. Country-level means and difference across cycles for Rapid guessing (OECD countries)

Country / Economy	Reading						Science						Mathematics					
	2018		2015		Difference		2018		2015		Difference		2018		2015		Difference	
	Mean	S.E.	Mean	S.E.	Diff.	S.E.	Mean	S.E.	Mean	S.E.	Diff.	S.E.	Mean	S.E.	Mean	S.E.	Diff.	S.E.
Australia	**0.03**	(0.002)	**0.03**	(0.002)	0.00	(0.002)	**0.04**	(0.001)	**0.03**	(0.001)	0.00	(0.002)	**0.02**	(0.001)	**0.02**	(0.001)	0.00	(0.002)
Austria	**0.03**	(0.003)	**0.01**	(0.002)	**0.02**	(0.003)	**0.03**	(0.002)	**0.01**	(0.001)	**0.02**	(0.002)	**0.01**	(0.002)	**0.01**	(0.001)	**0.01**	(0.002)
Belgium	**0.02**	(0.002)	**0.01**	(0.001)	**0.01**	(0.002)	**0.03**	(0.001)	**0.01**	(0.000)	**0.02**	(0.001)	**0.01**	(0.001)	**0.01**	(0.001)	**0.00**	(0.001)
Canada	**0.02**	(0.001)	**0.01**	(0.001)	**0.01**	(0.001)	**0.03**	(0.001)	**0.01**	(0.000)	**0.02**	(0.001)	**0.01**	(0.001)	**0.00**	(0.000)	**0.01**	(0.001)
Chile	**0.04**	(0.002)	**0.01**	(0.001)	**0.03**	(0.003)	**0.05**	(0.002)	**0.02**	(0.001)	**0.03**	(0.003)	**0.02**	(0.002)	**0.01**	(0.001)	**0.01**	(0.002)
Colombia	**0.06**	(0.004)	**0.02**	(0.001)	**0.04**	(0.005)	**0.06**	(0.004)	**0.02**	(0.001)	**0.04**	(0.004)	**0.02**	(0.003)	**0.01**	(0.001)	**0.01**	(0.003)
Costa Rica	**0.02**	(0.003)	**0.01**	(0.002)	**0.01**	(0.003)	**0.03**	(0.002)	**0.01**	(0.001)	**0.02**	(0.003)	**0.01**	(0.002)	**0.00**	(0.001)	**0.01**	(0.002)
Czech Republic	**0.03**	(0.003)	**0.02**	(0.001)	**0.01**	(0.003)	**0.03**	(0.002)	**0.02**	(0.001)	**0.01**	(0.002)	**0.01**	(0.001)	**0.01**	(0.001)	0.00	(0.002)
Denmark	**0.02**	(0.001)	**0.02**	(0.002)	**0.01**	(0.002)	**0.03**	(0.002)	**0.02**	(0.001)	**0.01**	(0.002)	**0.01**	(0.001)	**0.01**	(0.001)	0.00	(0.002)
Estonia	**0.02**	(0.001)	**0.01**	(0.001)	**0.01**	(0.001)	**0.02**	(0.001)	**0.01**	(0.001)	**0.01**	(0.001)	**0.01**	(0.001)	**0.01**	(0.001)	0.00	(0.001)
Finland	**0.02**	(0.001)	**0.02**	(0.001)	0.00	(0.002)	**0.02**	(0.001)	**0.01**	(0.001)	**0.01**	(0.001)	**0.01**	(0.001)	**0.01**	(0.001)	0.00	(0.001)
France	**0.04**	(0.002)	**0.02**	(0.001)	**0.02**	(0.002)	**0.04**	(0.002)	**0.02**	(0.001)	**0.03**	(0.002)	**0.02**	(0.001)	**0.01**	(0.001)	**0.01**	(0.002)
Germany	**0.03**	(0.003)	**0.10**	(0.004)	**-0.07**	(0.005)	**0.03**	(0.002)	**0.11**	(0.005)	**-0.07**	(0.005)	**0.01**	(0.001)	**0.06**	(0.004)	**-0.04**	(0.004)
Greece	**0.06**	(0.004)	**0.02**	(0.002)	**0.04**	(0.004)	**0.06**	(0.003)	**0.02**	(0.003)	**0.04**	(0.004)	**0.02**	(0.002)	**0.01**	(0.001)	**0.01**	(0.002)
Hungary	**0.03**	(0.002)	**0.01**	(0.002)	**0.02**	(0.003)	**0.03**	(0.002)	**0.01**	(0.001)	**0.02**	(0.002)	**0.01**	(0.002)	**0.01**	(0.001)	**0.01**	(0.002)
Iceland	**0.04**	(0.002)	**0.01**	(0.001)	**0.03**	(0.002)	**0.04**	(0.002)	**0.01**	(0.001)	**0.02**	(0.002)	**0.02**	(0.002)	**0.01**	(0.002)	**0.01**	(0.002)
Ireland	**0.02**	(0.001)	**0.03**	(0.002)	**-0.01**	(0.002)	**0.02**	(0.001)	**0.03**	(0.001)	**0.00**	(0.002)	**0.01**	(0.001)	**0.01**	(0.001)	**-0.01**	(0.002)
Israel	**0.07**	(0.005)	**0.03**	(0.002)	**0.04**	(0.005)	**0.07**	(0.004)	**0.03**	(0.001)	**0.04**	(0.004)	**0.04**	(0.004)	**0.01**	(0.001)	**0.03**	(0.004)
Italy	**0.05**	(0.003)	**0.01**	(0.001)	**0.04**	(0.003)	**0.05**	(0.003)	**0.01**	(0.001)	**0.04**	(0.003)	**0.02**	(0.001)	**0.01**	(0.001)	**0.01**	(0.002)
Japan	**0.03**	(0.002)	**0.01**	(0.001)	**0.02**	(0.002)	**0.02**	(0.001)	**0.01**	(0.000)	**0.02**	(0.001)	**0.01**	(0.001)	**0.00**	(0.001)	**0.00**	(0.001)
Korea	**0.04**	(0.003)	**0.02**	(0.002)	**0.02**	(0.004)	**0.04**	(0.003)	**0.02**	(0.001)	**0.02**	(0.003)	**0.03**	(0.003)	**0.01**	(0.001)	**0.02**	(0.003)
Latvia	**0.03**	(0.002)	**0.04**	(0.003)	**-0.01**	(0.003)	**0.03**	(0.002)	**0.04**	(0.002)	**-0.01**	(0.002)	**0.01**	(0.001)	**0.02**	(0.002)	**-0.01**	(0.002)

Country / Economy	Reading						Science						Mathematics					
	2018		2015		Difference		2018		2015		Difference		2018		2015		Difference	
	Mean	S.E.	Mean	S.E.	Diff.	S.E.	Mean	S.E.	Mean	S.E.	Diff.	S.E.	Mean	S.E.	Mean	S.E.	Diff.	S.E.
Lithuania	**0.02**	(0.001)	**0.02**	(0.002)	**0.01**	(0.002)	**0.03**	(0.002)	**0.02**	(0.001)	**0.01**	(0.002)	**0.01**	(0.001)	**0.01**	(0.001)	0.00	(0.002)
Luxembourg	**0.03**	(0.001)	**0.01**	(0.001)	**0.02**	(0.002)	**0.03**	(0.001)	**0.01**	(0.001)	**0.02**	(0.002)	**0.01**	(0.001)	**0.01**	(0.001)	**0.01**	(0.001)
Mexico	**0.01**	(0.002)	**0.02**	(0.002)	**-0.01**	(0.003)	**0.02**	(0.002)	**0.02**	(0.001)	0.00	(0.002)	**0.01**	(0.002)	**0.01**	(0.002)	0.00	(0.002)
Netherlands	**0.02**	(0.003)	**0.02**	(0.002)	**0.01**	(0.003)	**0.03**	(0.002)	**0.01**	(0.001)	**0.02**	(0.002)	**0.01**	(0.001)	**0.01**	(0.001)	0.00	(0.002)
New Zealand	**0.03**	(0.002)	**0.02**	(0.002)	**0.01**	(0.003)	**0.04**	(0.002)	**0.01**	(0.001)	**0.02**	(0.002)	**0.01**	(0.001)	**0.01**	(0.001)	**0.01**	(0.002)
Norway	**0.04**	(0.003)	**0.01**	(0.001)	**0.04**	(0.003)	**0.05**	(0.002)	**0.01**	(0.000)	**0.04**	(0.002)	**0.02**	(0.002)	**0.00**	(0.001)	**0.02**	(0.002)
Poland	**0.03**	(0.002)	**0.02**	(0.002)	**0.01**	(0.003)	**0.03**	(0.001)	**0.02**	(0.001)	**0.01**	(0.002)	**0.01**	(0.001)	**0.01**	(0.001)	**0.00**	(0.002)
Portugal	**0.03**	(0.002)	**0.02**	(0.002)	**0.01**	(0.003)	**0.03**	(0.002)	**0.02**	(0.001)	**0.01**	(0.002)	**0.01**	(0.001)	**0.02**	(0.001)	0.00	(0.002)
Slovak Republic	**0.04**	(0.002)	**0.01**	(0.001)	**0.04**	(0.002)	**0.05**	(0.002)	**0.01**	(0.000)	**0.04**	(0.002)	**0.02**	(0.002)	**0.00**	(0.000)	**0.01**	(0.002)
Slovenia	**0.02**	(0.001)	**0.08**	(0.004)	**-0.06**	(0.004)	**0.02**	(0.001)	**0.08**	(0.002)	**-0.06**	(0.002)	**0.01**	(0.001)	**0.05**	(0.002)	**-0.04**	(0.003)
Spain	**0.03**	(0.002)	**0.08**	(0.003)	**-0.05**	(0.004)	**0.04**	(0.002)	**0.07**	(0.002)	**-0.03**	(0.003)	**0.01**	(0.001)	**0.05**	(0.002)	**-0.03**	(0.002)
Sweden	**0.05**	(0.004)	**0.01**	(0.001)	**0.04**	(0.004)	**0.06**	(0.004)	**0.01**	(0.001)	**0.04**	(0.004)	**0.02**	(0.002)	**0.01**	(0.001)	**0.01**	(0.003)
Switzerland	**0.03**	(0.002)	**0.01**	(0.001)	**0.02**	(0.003)	**0.04**	(0.002)	**0.01**	(0.001)	**0.02**	(0.002)	**0.01**	(0.001)	**0.01**	(0.001)	**0.01**	(0.001)
Türkiye	**0.01**	(0.001)	**0.02**	(0.002)	0.00	(0.003)	**0.02**	(0.001)	**0.02**	(0.001)	**0.00**	(0.001)	**0.01**	(0.001)	**0.01**	(0.001)	**0.00**	(0.001)
United Kingdom	**0.04**	(0.002)	**0.01**	(0.001)	**0.03**	(0.003)	**0.05**	(0.003)	**0.01**	(0.001)	**0.04**	(0.003)	**0.02**	(0.002)	**0.01**	(0.001)	**0.02**	(0.002)
United States	**0.02**	(0.002)	**0.01**	(0.001)	**0.01**	(0.002)	**0.03**	(0.001)	**0.01**	(0.001)	**0.01**	(0.002)	**0.01**	(0.001)	**0.00**	(0.000)	**0.00**	(0.001)

Note: Means refer to the unstandardised measures, with higher values indicating higher levels of disengagement.

Table H.2. Country-level means and difference across cycles for Rapid guessing (Partner countries and economies)

Country / Economy	Reading						Science						Mathematics					
	2018		2015		Difference		2018		2015		Difference		2018		2015		Difference	
	Mean	S.E.	Mean	S.E.	Diff.	S.E.	Mean	S.E.	Mean	S.E.	Diff.	S.E.	Mean	S.E.	Mean	S.E.	Diff.	S.E.
Brazil	**0.07**	(0.003)	**0.01**	(0.001)	**0.06**	(0.003)	**0.07**	(0.003)	**0.01**	(0.001)	**0.06**	(0.003)	**0.03**	(0.002)	**0.01**	(0.001)	**0.03**	(0.002)
Bulgaria	**0.09**	(0.006)	**0.00**	(0.001)	**0.09**	(0.006)	**0.08**	(0.004)	**0.00**	(0.000)	**0.07**	(0.004)	**0.04**	(0.004)	**0.00**	(0.000)	**0.03**	(0.004)
Croatia	**0.03**	(0.003)	**0.01**	(0.001)	**0.01**	(0.003)	**0.03**	(0.002)	**0.01**	(0.001)	**0.02**	(0.003)	**0.02**	(0.002)	**0.01**	(0.001)	**0.01**	(0.002)
Dominican Republic	**0.14**	(0.006)	**0.01**	(0.001)	**0.14**	(0.006)	**0.14**	(0.006)	**0.01**	(0.001)	**0.13**	(0.006)	**0.06**	(0.003)	**0.00**	(0.001)	**0.05**	(0.003)
Hong Kong (China)	**0.03**	(0.002)	**0.01**	(0.001)	**0.02**	(0.002)	**0.03**	(0.002)	**0.01**	(0.001)	**0.03**	(0.002)	**0.01**	(0.001)	**0.00**	(0.000)	**0.01**	(0.001)
Macao (China)	**0.01**	(0.001)	**0.01**	(0.001)	0.00	(0.002)	**0.01**	(0.001)	**0.01**	(0.001)	**0.00**	(0.001)	**0.00**	(0.000)	**0.00**	(0.001)	**0.00**	(0.001)
Montenegro	**0.07**	(0.002)	**0.02**	(0.002)	**0.05**	(0.003)	**0.06**	(0.002)	**0.02**	(0.001)	**0.05**	(0.002)	**0.03**	(0.002)	**0.01**	(0.001)	**0.02**	(0.002)
Peru	**0.02**	(0.002)	**0.01**	(0.001)	**0.02**	(0.002)	**0.03**	(0.001)	**0.01**	(0.001)	**0.02**	(0.001)	**0.01**	(0.001)	**0.00**	(0.001)	**0.00**	(0.001)
Qatar	**0.12**	(0.002)	**0.02**	(0.001)	**0.10**	(0.002)	**0.11**	(0.002)	**0.02**	(0.001)	**0.09**	(0.002)	**0.06**	(0.002)	**0.01**	(0.001)	**0.05**	(0.002)
Singapore	**0.02**	(0.001)	**0.02**	(0.001)	0.00	(0.002)	**0.03**	(0.002)	**0.01**	(0.001)	**0.02**	(0.002)	**0.01**	(0.001)	**0.01**	(0.001)	0.00	(0.002)
Chinese Taipei	**0.03**	(0.002)	**0.01**	(0.001)	**0.02**	(0.002)	**0.04**	(0.002)	**0.02**	(0.001)	**0.02**	(0.002)	**0.02**	(0.002)	**0.01**	(0.001)	**0.01**	(0.002)
Thailand	**0.02**	(0.003)	**0.01**	(0.002)	**0.01**	(0.003)	**0.03**	(0.003)	**0.01**	(0.001)	**0.02**	(0.003)	**0.01**	(0.002)	**0.01**	(0.001)	**0.01**	(0.003)
United Arab Emirates	**0.05**	(0.002)	**0.01**	(0.001)	**0.04**	(0.002)	**0.06**	(0.002)	**0.01**	(0.000)	**0.04**	(0.002)	**0.03**	(0.001)	**0.01**	(0.001)	**0.02**	(0.001)
Uruguay	**0.06**	(0.004)	**0.01**	(0.001)	**0.04**	(0.004)	**0.05**	(0.003)	**0.01**	(0.001)	**0.04**	(0.003)	**0.02**	(0.002)	**0.01**	(0.001)	**0.02**	(0.002)

Notes: Means refer to the unstandardised measures, with higher values indicating higher levels of disengagement. Only countries/economies are displayed in which self-reported effort was administered in both cycles.