Assessing Problem-Solving Skills in PISA 2012

This chapter introduces the PISA 2012 assessment of problem solving. It provides the rationale for assessing problem-solving competence in PISA, and introduces the innovative features of the 2012 assessment. The framework for the assessment is presented, and sample items are discussed.
In Daniel Defoe’s novel, Robinson Crusoe is stranded on a desert island. He first needs to secure food for himself. To solve this problem, he re-invents agriculture and tames a flock of wild goats. Then, he returns to his true longing: “My desire to venture over for the main[land] increased, rather than decreased, as the means for it seemed impossible. This at length put me upon thinking whether it was not possible to make myself a canoe […], even without tools, […] of the trunk of a great tree. This I not only thought possible, but easy” (Defoe, 1919).

Problems are situations with no obvious solution, and solving problems requires thinking and learning in action. Problem solving “involves initiating, usually on the basis of hunches or feelings, experimental interactions with the environment to clarify the nature of a problem and potential solutions”, so that the problem-solver “can learn more […] about the nature of the problem and the effectiveness of their strategies”, “modify their behaviour and launch a further round of experimental interactions with the environment” (Raven, 2000, p. 54). (Robinson Crusoe’s first strategy to escape from his island in a canoe fails, for, as he explains, “my thoughts were so intent upon my voyage over the sea in [the canoe], that I never once considered how I should get it off the land”.)

Just like Robinson Crusoe, we solve small problems every day: “My mobile phone has stopped working; how do I tell my friends that I’m running late for our appointment?”; “This meeting room is so cold; are these the switches to control the air conditioning?”; “I don’t speak the local language, and my connecting flight leaves from a different airport in the same city. I just hope I can get there in time.”

In modern societies, all of life is problem solving. Changes in society, the environment and in technology mean that the content of applicable knowledge evolves rapidly. Today’s 15-year-olds are the Robinson Crusoes of a future that remains largely unknown to us. Adapting, learning, daring to try out new things, and always being ready to learn from mistakes are among the keys to resilience and success in an unpredictable world.

This chapter begins with a discussion of the rationale for including a separate assessment of problem solving in PISA. It then introduces what is new and distinctive about the PISA 2012 approach to assessing problem solving, and describes the main dimensions covered in the problem-solving framework. The chapter concludes by presenting the test interface and sample items from the PISA computer-based assessment of problem solving.

**WHY PISA ASSESSES PROBLEM-SOLVING COMPETENCE**

Today’s workplaces demand people who can solve non-routine problems. Few workers, whether in manual or knowledge-based occupations, use repetitive actions to perform their job tasks. The Survey of Adult Skills (PIAAC), for instance, measured how often workers are faced with a new or difficult situation in their jobs that requires some thinking before taking action (OECD, 2013a). On average across countries, a large majority of workers are confronted at least once per week in their job with simple problems (those requiring less than 30 minutes to find a solution). Meanwhile, one in ten workers is confronted every day with more complex problems that require at least 30 minutes to find a good solution. Complex problem-solving skills are particularly in demand in fast-growing, highly skilled managerial, professional and technical occupations.

One possible explanation for this shift to non-routine tasks in the workplace is that, as computers and computerised machines were introduced in greater numbers, workers were needed less often to perform routine manual or analytical tasks. Instead, they were required to deal with the unexpected and the unfamiliar, and to bring the best out of the machines and computers working alongside them (Autor, Levy and Murnane, 2003). There is clear evidence of this change in the demand for skills in Germany, Japan and the United States (Box V.1.1 and Figure V.1.1).

Acknowledging these changes, the emphasis in education is shifting too, from equipping students with highly codified, routine skills to empowering them to confront and overcome complex, non-routine cognitive challenges. Indeed, the skills that are easiest to teach and test are also the skills that are easiest to digitise, automate and outsource. For students to be prepared for tomorrow’s world, they need more than mastery of a repertoire of facts and procedures; students need to become lifelong learners who can handle unfamiliar situations where the effect of their interventions is not predictable. When asked to solve problems for which they have no ready-made strategy, they need to be able to think flexibly and creatively about how to overcome the barriers that stand in the way of a solution.

Non vitae, sed scholae discimus

[Too often,] we don’t learn for life, but only for the lecture room

Seneca, Ad Lucillium, c. 65 AD
Box V.1.1. **Long-term trends in the demand for problem-solving skills**

Trends in the demand for skills can be inferred from aggregate measures of workers’ job requirements, repeated over time. Figure V.1.1 presents the observed evolution of job requirements in three major OECD countries: Germany, Japan and the United States. Across all three countries, there has been a marked increase in the demand for problem-solving skills.

According to Autor, Levy and Murnane (2003), job requirements can be classified into five major skill categories. A first distinction is between “routine” and “non-routine” tasks and skills. “Routine” skills correspond to tasks that “require methodical repetition of an unwavering procedure” (p. 1283), i.e. those tasks in which machines and computers can fairly easily replace human beings. They can be cognitive (such as data entry) or manual (such as repetitive production). “Non-routine” skills correspond to tasks that require tacit knowledge and are only imperfectly described in terms of a set of rules.

A further distinction, within non-routine skills, is between “manual” and “abstract” skills. Manual non-routine tasks, such as preparing a meal, demand situational adaptability, visual and language recognition, and interaction with other people. They are difficult to automate, but from the human perspective, they are straightforward, requiring primarily abilities that are hardwired into humans’ evolutionary endowments. Abstract tasks are based on the processing of information and require problem-solving skills, intuition, persuasion and creativity. Among abstract skills, there are “analytic” and “interpersonal” skills: “interpersonal” tasks (such as managing teams or persuading potential buyers) require complex interpersonal communication, while “analytic” tasks require the transformation of data and information.
Problem-solving competence is an essential component of the skills required to perform interpersonal and non-routine analytic tasks successfully. In both kinds of tasks, workers need to think about how to engage with the situation, monitor the effect of their actions systematically, and adjust to feedback.

In Germany, a representative sample of workers has consistently reported on job requirements over more than 20 years, providing direct evidence of an increase in the use of non-routine analytic and interactive skills in the workplace during the 1980s and 1990s (Spitz-Oener, 2006). This increase has been accompanied by declines in the importance of routine skills, both analytic (such as skills needed for bookkeeping) and manual (such as sorting).

In the United States and Japan, the evolution of aggregate skill requirements has been estimated by matching job titles reported to the national population census with precise job descriptions in the dictionary of occupational titles, for the United States (Autor, Levy and Murnane, 2003; Autor and Price, 2013), or in the career matrix constructed by the Institute for Labour Policy and Training in Japan (Ikenaga and Kambayashi, 2010). Changes in the occupational shares for precisely defined occupations can then be translated into changes in the economy’s skill requirements. This methodology has yielded strikingly similar results as found in Germany, over a longer period of time, i.e. since 1960.

While problem-solving skills are increasingly needed in today’s economies, the ability to adapt to new circumstances, learn throughout life, and turn knowledge into action has always been important for full participation in society. The best educators have always aimed to foster the skills needed to perform non-routine tasks, i.e. to teach for life, not for school.

Recent evidence confirms that the generic skills examined in a problem-solving assessment such as PISA are strongly associated with academic success and are distinct from reasoning or intelligence, as traditionally measured (Wüstemberg et al., 2012; Greiff et al., 2013a; Funke and Frensch, 2007). In addition, other research strongly supports the view that good teachers and schools can develop students’ overall problem-solving skills through and in addition to their competence in regular curricular subjects (Csapó and Funke, forthcoming).

Yet all too often teachers find that while their students may excel on routine exercises (those that they have already seen and practiced), they fail to solve problems that are unlike those they have previously encountered. Clearly, mastering the simple steps that are required to reach a solution is not enough. Students need to be able to know not only what to do, but also when to do it; and they need to feel motivated and interested. Mayer (1998) summarises these three components of successful problem solving in all domains as “skill”, “metaskill” and “will”.

The problem-solving assessment in PISA 2012 focuses on students’ general reasoning skills, their ability to regulate problem-solving processes, and their willingness to do so, by confronting students with problems that do not require expert knowledge to solve. Individual problem solving was assessed as a separate domain for the first time in 2003 (OECD, 2005). The advances in our understanding of problem solving since then and the opportunities afforded by computers to improve the assessment of problem-solving skills led to the inclusion of problem solving as a core component of the PISA 2012 assessment.1

The regular assessments of mathematics, reading and science in PISA all include problem-solving tasks that assess students’ ability to use their curricular knowledge to meet real-life challenges. Indeed, problem-solving competence need not be developed independently of expertise in curricular subjects; in fact, the literature on the development of general cognitive abilities suggests that content-based methods can be equally effective and may be preferable: “If you teach the specifics with abstraction in mind, the general is learned, but if you try to teach the general directly, the specifics are often not learned” (Adley et al., 2007, p. 92).

While schools are not the only environment in which problem-solving competence is nurtured, high-quality education, in a wide range of subjects, certainly helps to develop these skills. Progressive teaching methods, like problem-based learning, inquiry-based learning, and individual and group project work, can be used to foster deep understanding and prepare students to apply their knowledge in novel situations. Good teaching promotes self-regulated learning and metacognition – particularly knowledge about when and how to use certain strategies for learning or for problem solving – and develops cognitive dispositions that underpin problem solving. It prepares students to reason effectively in unfamiliar situations, and to fill gaps in their knowledge by observing, exploring and interacting with unknown systems.
All teachers can create opportunities to develop problem-solving competence. For instance, thinking habits, such as careful observation, awareness about one’s working process, or critical self-evaluation, can be instilled in students as they learn techniques in the visual arts (Winner et al., 2013; see Box V.5.5) – and indeed, in any other subject in the school curriculum. Because the skills and dispositions that underpin successful problem solving in real life are not specific to particular subjects, students who learn to master them in several curricular contexts will be better equipped to use them outside of school as well.

Thus, by measuring 15-year-olds’ problem-solving skills, PISA provides evidence about the comparative success of education systems in equipping students for success in life, evidence that can, in turn, inform education policies and practices.

**THE PISA 2012 APPROACH TO ASSESSING STUDENT PERFORMANCE IN PROBLEM SOLVING**

The problem-solving assessment in PISA 2012 focuses on general cognitive processes involved in problem solving, rather than on the ability to solve problems in particular school subjects. Given the advances in understanding the cognitive processes involved in problem solving and the possibility of using computer-based simulated scenarios, the assessment also assigns a central place to so-called interactive problems.

**A focus on general cognitive processes involved in solving problems**

Research findings suggest that outside of artificial laboratory conditions, the situation in which a problem is embedded influences the strategies used to solve it (Kotovsky, Hayes and Simon, 1985; Funke, 1992). In real life, highly proficient problem-solvers in one context may act as novices when confronted with problems outside of their field of expertise.

In the context of a particular subject, trade or occupation, experts will use domain-specific knowledge and strategies to solve the problems. Meanwhile, those who solve problems efficiently, even when they arise outside of their field of expertise, have mastered general reasoning skills, can apply those skills where appropriate, and are motivated to engage with unfamiliar problems.

A glimpse at some of the names of problem-solving units included in the PISA assessment reveals the typical contexts included in the assessment: technology devices (e.g. REMOTE CONTROL, CLOCK, LIGHTS), unfamiliar spaces (e.g. TRAFFIC, LOST), food or drink (e.g. VITAMINS, DRINK MACHINE), etc. These contexts refer to situations that students may encounter outside of school as part of their everyday experience.

While including authentic scenarios related to real-life problems, the PISA 2012 problem-solving assessment avoids the need for specific, curricular knowledge as much as possible. Texts are short and use plain language. If arithmetic operations are required, calculators are embedded in the scenario. In contrast, when problem-solving tasks are incorporated in the assessment of the regular PISA domains of mathematics, reading and science, expert knowledge in these areas is needed in order to reach a solution.

By using authentic problem situations, the assessment also reduces the influence of affective factors related to school, or to specific subjects, on results. The student’s familiarity with the context may still influence how he or she approaches the problem. Because the assessment tasks are embedded in real-life settings, in practice some students may be more familiar than others with the concrete contexts. However, since a wide range of contexts is included in the different assessment units, the degree of familiarity with the setting will vary, so that prior knowledge will not systematically influence performance. In addition, applying prior knowledge is never sufficient for solving new problems, even in familiar situations.

**The centrality of interactive problem solving**

In most problems that students practice in class or when studying for an exam, the information needed to solve the problem is provided at the outset. By contrast, solving real-life problems often requires identifying the pieces of information available in the environment/context that would be most useful for solving the problem.

Problems that require students to uncover useful information by exploring the problem situation are called interactive problems. These kinds of problems are encountered when using unfamiliar everyday devices, such as a new mobile phone, home appliance or vending machine. Outside of technological contexts, similar situations also arise in social interactions and in other settings as varied as cultivating plants or raising animals. A majority of PISA 2012 problem-solving tasks correspond to interactive problems. The prevalence of interactive problems in the PISA 2012 assessment reflects their importance in the real world.
The inclusion of interactive tasks, made possible by computer delivery, represents the main innovation over the PISA 2003 assessment of problem solving. PISA 2012 therefore provides a broader measure of problem-solving competency than previous assessments of problem solving.

The PISA definition of problem-solving competence

PISA 2012 defines problem-solving competence as:

...an individual’s capacity to engage in cognitive processing to understand and resolve problem situations where a method of solution is not immediately obvious. It includes the willingness to engage with such situations in order to achieve one’s potential as a constructive and reflective citizen.

The PISA 2012 framework publication (OECD, 2013b) discusses the definition in full. Among the key elements:

... an individual’s capacity to engage in cognitive processing to understand and resolve problem situations...

Problem solving begins with recognising that a problem situation exists and establishing an understanding of the nature of the situation. It requires the solver to identify the specific problem(s) to be solved, plan and carry out a solution, and monitor and evaluate progress throughout the activity.

The verbs engage, understand and resolve underline that, in addition to the explicit responses to items, the assessment measures individuals’ progress towards solving a problem, including the strategies they employ. Where appropriate, these strategies are tracked through behavioural data captured by the computer.

... where a method of solution is not immediately obvious...

This part of the definition corresponds to the definition of “problem” as a situation in which the goal cannot be achieved by merely applying previously learned procedures (Mayer, 1990). The PISA assessment of problem solving is only concerned with such non-routine tasks.

In many real-life situations, the same task may be considered a novel problem by some and a routine problem by others. With learning and practice, some activities that were initially experienced as problem solving may become routine activities. The problems included in the PISA assessment of problem solving involve tasks that are non-routine for 15-year-old students. Although some students may be familiar with the context or the goal of a problem situation that refers to a plausible real-world scenario, the particular problem faced is novel and the ways of achieving the goal are not immediately obvious.

For example, consider the problem of determining whether a lamp is not working because a) the switch is malfunctioning, b) there is no power, or c) the light bulb needs to be changed. Although the situation might be familiar to many 15-year-olds, few students, if any, have had the opportunity to develop expertise in this class of problems, and the unique design of a test unit around this problem situation makes sure that at least some adaptation of ready-made strategies is needed.

Even in non-routine problems, however, the knowledge of general strategies, including those learned at school, can be of help. The lamp problem described above is a case in point. As in many problems where the solver needs to develop an understanding of cause-effect relationships, an effective approach is to “vary one thing at a time”. This strategy is at the heart of the experimental method in the natural sciences and is taught as such in school curricula throughout the world. Several problem-solving units included in the PISA assessment indirectly require students to apply a particular strategy in non-curricular contexts, without being prompted to do so.

... it includes the willingness to engage with such situations...

The last sentence of the definition underscores that the use of knowledge and skills to solve a problem depends on motivational and affective factors as well (Mayer, 1998; Funke, 2010). Students’ willingness to engage with novel situations is an integral part of problem-solving competence. Motivational and affective factors are a distinct focus of the background questionnaire, which uses students’ answers to measure their perseverance (whether they agree or not with the statement “When confronted with a problem, I give up easily”, and other similar statements) and openness to problem solving (“I like to solve complex problems”).
THE PISA 2012 FRAMEWORK FOR ASSESSING PROBLEM-SOLVING COMPETENCE

The PISA framework for assessing problem-solving competence guided the development of the assessment and sets the parameters for reporting results. The framework identifies three distinct aspects: the nature of the problem situation, the problem-solving processes involved in each task, and the problem context. The main elements of the problem-solving framework are summarised in Figure V.1.2.

The nature of the problem situation is determined by whether the information disclosed to the student at the outset is sufficient to solve the problem (static problems), or whether interaction with the problem situation is a necessary part of the solving activity (interactive problems). Examples of interactive problems include problems commonly faced when using unfamiliar devices, such as a new mobile phone or a ticket-vending machine.

For the purpose of the PISA assessment, the cognitive processes involved in problem solving are grouped into four problem-solving processes:

- **Exploring and understanding.** This involves exploring the problem situation by observing it, interacting with it, searching for information and finding limitations or obstacles; and demonstrating understanding of the information given and the information discovered while interacting with the problem situation.

- **Representing and formulating.** This involves using tables, graphs, symbols or words to represent aspects of the problem situation; and formulating hypotheses about the relevant factors in a problem and the relationships between them, to build a coherent mental representation of the problem situation.

- **Planning and executing.** This involves devising a plan or strategy to solve the problem, and executing it. It may involve clarifying the overall goal, setting subgoals, etc.

- **Monitoring and reflecting.** This involves monitoring progress, reacting to feedback, and reflecting on the solution, the information provided with the problem, or the strategy adopted.

No assumption is made that the processes involved in solving a particular problem are sequential or that all of the processes listed are involved in solving a particular problem. As individuals confront, represent and solve problems, they may move to a solution in a way that transcends the boundaries of a linear, step-by-step model. Nevertheless, single items were intended to have one of these processes as their main focus.

Although reasoning skills were not explicitly used to organise the domain, each of the problem-solving processes draws upon one or more of them. In understanding a problem situation, the solvers may need to distinguish between facts and
opinion; in formulating a solution, they may need to identify relationships between variables; in selecting a strategy, they may need to consider cause and effect; and, in reflecting on results, they may need to critically evaluate assumptions and alternative solutions. Deductive, inductive, analogical, combinatorial, and other types of reasoning are embedded within problem-solving tasks in PISA. It is important to note that these types of thinking can be taught and honed in classroom instruction (e.g. Adey et al., 2007; Klauer and Phye, 2008).

The problem context is classified according to two dimensions: technology or non-technology, and personal or social. Problems in technology settings involve a technological device, such as a digital clock, an air conditioner, or a ticket machine; problems in non-technology settings do not, and include problems such as task scheduling or decision making. Problems with a personal focus refer to situations involving only the student, the student’s family or close peers; problems with a social focus relate to situations encountered more broadly in the community or society in general.

Items were developed to measure how well students perform when the various problem-solving processes are exercised within the two different types of problem situations across a range of contexts. Each of these key aspects is discussed and illustrated in Chapter 3.

THE DESIGN AND DELIVERY OF THE PISA 2012 COMPUTER-BASED ASSESSMENT OF PROBLEM SOLVING

The development of items for the assessment

As in all other domains, the items for the PISA 2012 problem-solving assessment came from two sources: the PISA Consortium and national submissions. The problem solving expert group that developed the PISA 2012 framework reviewed all materials to ensure that they reflected the defined construct of problem-solving competence. The items were then reviewed by national centres and field tested. If the national review indicated significant concern that an item would advantage a particular country or language group, it was not considered for inclusion in the main assessment. The procedures to ensure that no group would be consistently advantaged (or disadvantaged) by a particular item are described in greater detail in the PISA 2012 Technical Report (OECD, forthcoming).

A variety of response formats were used, including many that were only possible because the assessment was delivered by computer, such as the use of drop-down menus for selected response formats, or constructed responses coded automatically.

As usual in PISA, items are arranged in units grouped around a common stimulus. The survey included 16 units, with a total of 42 items. Sample units from the PISA assessment of problem solving are introduced and described at the end of this chapter.

The structure and delivery of the assessment

In the 28 OECD countries and 16 partner countries and economies that participated in the assessment of problem solving, the survey was conducted after the paper-based assessment of mathematics, reading and science. In countries that also assessed mathematics and reading on computers, these computer-based tests were administered at the same time as the problem-solving assessment. The 16 units of the problem-solving assessment were grouped into four clusters, each of which was designed to be completed in 20 minutes. Each student assessed was given either one or two clusters, depending on whether the student was also participating in the computer-based assessment of mathematics or reading. In all cases, the total time allocated to computer-based tests was 40 minutes.

The appearance of the test interface was consistent across items (see Figure V.1.3 for an example). For each item the stimulus material appeared in the top part of the screen. The item appeared in the lower part of the screen, and was separated visually from the stimulus by borders. The points at which the screen was divided varied from item to item so that scrolling was never required.

Test units within clusters and single items within units were delivered in a fixed order, with no possibility of returning to a previous item once students had begun the next item. Each test item, with its associated stimulus material, occupied a single computer screen. Students were asked to confirm that they wanted to proceed to the next item when they pressed the next item icon (arrow) in the bottom right corner of the test interface.
The opportunities afforded by computer delivery

PISA 2012 marks the second time that individual problem-solving competence was assessed in PISA. In 2003, a paper and pencil test of cross-disciplinary problem solving was part of the assessment (OECD, 2005). In PISA 2012, computer delivery was fundamental to the conception of problem solving. A paper-and-pencil assessment of problem solving could not have measured the same construct. The inclusion of interactive problems, in which students need to explore the (simulated) environment and gather feedback on the effect of their interventions in order to obtain all the information needed to solve a problem, was only possible by asking students to use a computer to complete the assessment.

In addition, information about how students interact with the material as they progressed through the assessment was stored on the computer. This information includes the types of actions a student completes (e.g. mouse click, drag and drop, keystrokes), the frequency of interaction between the student and the material, the sequence of actions, the state of the system at any given point, and the timing of specific interactions.

The computer delivery made it possible to include authentic response formats, where the observed behaviour corresponds to the answer. This is a major step towards evaluating authentic problem-solving performance. For instance, Question 1 from the unit TICKETS asks students to use a machine that they have never seen before to buy a ticket (Figure V.1.3); students earn credit if they succeed in buying the ticket. Students do not need to describe the process in a text or drawing field, or by ticking boxes. Various selected response formats, such as drop-down menus, were also included that would not have been possible in a paper-based test.

In several items the score reflects not only the explicit response given by students, but also the sequence of actions that they perform before giving the response. For example, in a hypothetical item that required students to troubleshoot a malfunctioning device, where students would need to explore the device in order to uncover information, students
would not get credit for selecting the broken element from a number of given possibilities unless the data logged by the computer indicated that the student had taken the necessary steps to rule out other plausible alternatives. One of the innovative features of the problem-solving assessment is that information contained in log files about the sequence of actions performed by students was used to inform scoring of items where appropriate. For example, when it could be established that students had guessed an answer, they received no credit for that answer.

Given that the assessment was delivered on computers, familiarity with information and communication technologies (ICT) may have influenced students’ performance. The ICT competence needed to navigate the test interface was limited to such basic skills as using a keyboard, a mouse or a touchpad, clicking radio buttons, dragging-and-dropping, scrolling and using pull-down menus and hyperlinks. In a further attempt to remove any advantage to students who were more familiar with computers, all students completed, before the assessment, a practice unit that contained examples of each of the response formats required.

**PROBLEM-SOLVING TASKS**

**General characteristics of static and interactive problem-solving tasks**

As in PISA 2003, static tasks include decision-making problems, where the student has to choose among alternatives under constraints, and system-analysis problems, where the student needs to identify relationships between parts of a system. The unit TRAFFIC is an example of a decision-making problem, and the unit ROBOT CLEANER is an example of a system-analysis problem (see the section on sample tasks below for more details on each unit).

In general, the five units with static items present analytical problems similar to those included in the PISA 2003 assessment of problem solving. However, since these items were delivered on a computer in 2012, PISA used new formats for the stimulus information (such as animations; see the unit ROBOT CLEANER) and new response formats (such as drag-and-drop).

Most interactive units included in the PISA 2012 assessment of problem solving belong to one of two classes of problems studied in the literature, “MicroDYN” systems and “finite-state automata”. In both cases, exploration and control of an unknown system are the two main tasks for the student. The single exception is a resource-allocation problem, in which experimental interaction with the test scenario is needed to uncover important information about the available resources.

Four units are MicroDYN units, based on small dynamic systems of causal relationships (Greiff et al., 2013b; Wüstenberg et al., 2012). The unit CLIMATE CONTROL provides an illustration. MicroDYN units share a common structure. They consist of a system of causal relations involving only a few variables that have to be explored and controlled in order to reach assigned goal states. In the first, “knowledge-generation” phase, the student has to control up to three input variables; a graph illustrates the effect of inputs on up to three output variables. Students typically have to demonstrate rule knowledge after this first phase. Students are then asked to control the system to reach a certain target by choosing the appropriate input levels. MicroDYN units vary in the way inputs and outputs are connected in a system, in the number of variables that the system comprises, and in the fictitious scenario in which interactions with the variables take place.

Six interactive units are based on finite-state automata (Buchner and Funke, 1993; Funke, 2001), including the unit TICKETS. The field trial unit MP3 PLAYER also belongs to this group. In contrast to MicroDYN units, the outcome of an intervention is not represented by a quantity, but by a new state of the system. Many of these units are based on everyday technological devices, and the behaviour of the device depends on both the current state and on the input command received from the user. The context need not be technological, however; a simulated navigation task, where students need to orient themselves by exploring an unfamiliar neighbourhood, is similar in form. What students see in the next step depends both on where they are and what action they take.

The distinctive characteristic of finite-state automata is that there are only a finite number of possible states (not all of which are known at the outset), and a limited number of input commands (whose effect may or may not be transparent at the outset). The effect of the interventions may, or may not, depend on the current state of the system. The amount of relevant information that needs to be discovered, the number of possible actions, and the number of possible states all contribute to the level of difficulty of the item.

In these problems, students typically need to explore the system or device in order to understand the effect of their interventions, explain the functioning of the device, bring the device into some desired state, or propose improvements to the device.
Sample tasks from the PISA 2012 problem-solving assessment

Items from one unit included in the PISA 2012 field trial, and from four units that were included in the PISA 2012 main survey, are described below. For each unit, a screenshot of the stimulus information is provided, together with a brief description of the context of the unit. This is followed by a screenshot and description of each item from that unit. The test units described below are also available for viewing on the web at http://cbasq.acer.edu.au. The interactive nature of the units MP3 PLAYER, CLIMATE CONTROL and TICKET MACHINE can be best appreciated by trying to solve the items.

Sample unit 1: MP3 PLAYER (field trial)

In the unit MP3 PLAYER, students are told that they have been given an MP3 player by a friend. They do not know how it works and must interact with it to find out, so the nature of the problem situation for each item in this unit is interactive. Since the focus of the unit is on discovering the rules that govern a device intended for use by an individual, the context of each item in the unit is technology and personal.

MP3 PLAYER: Item 1

Question 1: MP3 PLAYER CP043Q03
The bottom row of the MP3 player shows the settings that you have chosen. Decide whether each of the following statements about the MP3 player is true or false.
Select "True" or "False" for each statement to show your answer.

<table>
<thead>
<tr>
<th>Statement</th>
<th>True</th>
<th>False</th>
</tr>
</thead>
<tbody>
<tr>
<td>You need to use the middle button ( ) to change the type of music.</td>
<td>☑</td>
<td>☐</td>
</tr>
<tr>
<td>You have to set the volume before you can set the bass level.</td>
<td>☑</td>
<td>☐</td>
</tr>
<tr>
<td>Once you have increased the volume, you can only decrease it if you change the type of music you are listening to.</td>
<td>☑</td>
<td>☐</td>
</tr>
</tbody>
</table>
In the first item in the unit, students are given a series of statements about how the system works and are asked to identify whether the statements are true or false. The statements offer scaffolding for students to explore the system. The problem-solving process for this item is exploring and understanding, and the exploration is guided but unrestricted. A “Reset” button is available that allows students to return the player to its initial state at any time and start their exploration again if desired. There is no restriction on the number of times this can be done. In the field trial this was a somewhat harder-than-average item, with 38% of students gaining full credit (True, False, False), due probably to the requirement that all three answers must be correct and the degree to which information has to be uncovered (no information is known about the system at the outset and so all knowledge of the rules of the system must come from interacting with it). Partial credit was not available for this item.

**MP3 PLAYER: Item 2**

**Question 2: MP3 PLAYER CP043Q02**

Set the MP3 player to Rock, Volume 4, Bass 2.

Do this using as few clicks as possible. There is no RESET button.

The second item in the unit is classified as planning and executing. In this item, students must plan how to achieve a given goal and then execute this plan. Of interest for this partial-credit item is that process information captured by the computer (in this case, how many steps the student takes to successfully reach the goal state) contributes to the score. The task is to be completed using as few clicks as possible and the option of returning the machine to its initial state by pressing the “Reset” button is not available. If the number of clicks used (no more than 13) indicates that students have been efficient in reaching the goal they receive full credit; but if they reach the goal in a less-efficient manner (more than 13 clicks), they only receive partial credit. The requirement for efficiency made it more difficult to earn full credit for this item, though it was fairly easy to earn at least partial credit. In the field trial, about 39% of students received full credit and about 33% received partial credit.

**MP3 PLAYER: Item 3**

**Question 3: MP3 PLAYER CP043Q01**

Shown below are four pictures of the MP3 player's screen. Three of the screens cannot happen if the MP3 player is working properly. The remaining screen shows the MP3 player when it is working properly. Which screen shows the MP3 player working properly?
The third item in the unit is classified as representing and formulating since it requires students to form a mental representation of the way the whole system works in order to identify which of four given pictures shows the MP3 player when it is working properly. Returning the player to its initial state, which was possible in the first item, but absent in the second item of the unit, is again possible, so the student may interact with the system as much or as little as needed. Partial credit was not available for this item. In the field trial it was as difficult as the first item in the unit, with 39% of students selecting the correct response (the second option from the left).

**MP3 PLAYER: Item 4**

Question 4: MP3 PLAYER CP043Q04

Describe how you could change the way the MP3 player works so that there is no need to have the bottom button ( ). You must still be able to change the type of music, and increase or decrease the volume and the bass level.

The final item in this unit is classified as monitoring and reflecting, and asks students to reconceptualise the way the device works. This item is a constructed-response item and requires expert scoring. Full-credit answers are those that suggest how the MP3 player might operate with only two buttons instead of the original three. There is no single correct answer. Students may think creatively in devising a solution, but the most obvious solution is to suggest changing the way the top button works so that once you reach the right side of the display, one more click takes you back to the left of the display. In the field trial, this was by far the hardest item in the unit, likely because of the requirement of providing a constructed response and the item’s degree of abstraction: students must imagine a hypothetical scenario and link it to their mental representation of how the system currently works, in order to describe a possible alternative functioning. Only 25% of students earned credit; partial credit was not available for this item.

**Sample unit 2: CLIMATE CONTROL**

CLIMATE CONTROL: Stimulus information

You have no instructions for your new air conditioner. You need to work out how to use it.

You can change the top, central and bottom controls on the left by using the sliders ( ). The initial setting for each control is indicated by .

By clicking APPLY, you will see any changes in the temperature and humidity of the room in the temperature and humidity graphs. The box to the left of each graph shows the current level of temperature or humidity.
In the unit CLIMATE CONTROL, students are told that they have a new air conditioner but no instructions for it. Students can use three controls (sliders) to vary temperature and humidity levels, but first they need to understand which control does what. A measure of temperature and humidity in the room appears in the top-right part of the screen, both in numeric and in graphical form. All items in this unit present an interactive problem situation, with context classified as personal and technological.

The unit CLIMATE CONTROL is a typical MicroDYN unit, with a first “knowledge-generation” task and a second “knowledge-application” task. Knowledge generation in the MicroDYN environment requires students to carefully monitor the effects of their interventions. The increase in the level of an input variable leads either to an increase, a decrease, a mixed effect (increase and decrease for different variables), or to no effect in one or more output variables.

**CLIMATE CONTROL: Item 1**

**Question 1: CLIMATE CONTROL CP025Q01**
Find whether each control influences temperature and humidity by changing the sliders. You can start again by clicking RESET.
Draw lines in the diagram on the right to show what each control influences. To draw a line, click on a control and then click on either Temperature or Humidity. You can remove any line by clicking on it.

In the first item in the unit, students are invited to change the sliders to find out whether each control influences the temperature or the humidity level. The problem-solving process for this item is representing and formulating: the student must experiment to determine which controls have an impact on temperature and which on humidity, then represent the causal relations by drawing arrows between the three controls and the two outputs (temperature and humidity). There is no restriction on the number of rounds of exploration that the student is allowed. Full credit for this question requires that the causal diagram is correctly completed. Partial credit for this question is given if the student explores the relationships among variables efficiently, by varying only one input at a time, but fails to correctly represent them in a diagram.

**CLIMATE CONTROL: Item 2**

**Question 2: CLIMATE CONTROL CP025Q02**
The correct relationship between the three controls, Temperature and Humidity is shown on the right.
Use the controls to set the temperature and humidity to the target levels. **Do this in a maximum of four steps.** The target levels are shown by the red bands across the Temperature and Humidity graphs. The range of values for each target level is 18-20 and is shown to the left of each red band.
**You can only click APPLY four times and there is no RESET button.**
The second item in the unit asks students to apply their new knowledge of how the air conditioner works to set temperature and humidity at specified target levels (lower than the initial state). This is a planning and executing item. To ensure that no further exploration is needed beyond the one conducted in the previous item, a diagram shows how the controls are related to temperature and humidity levels (students could not return to any previous item during the test). Because only four rounds of manipulation are permitted, students need to plan a few steps ahead and use a systematic, if simple, strategy to succeed in this task. Nevertheless, the target levels of temperature and humidity provided can be reached in several ways within four steps – the minimum number of steps needed is two – and a mistake can often be corrected, if immediate remedial action is taken. A possible strategy, for instance, is to set separate subgoals and to focus on temperature and humidity in successive steps. If the student is able to bring temperature and humidity both closer to their target levels within the four rounds of manipulation permitted, but does not reach the target for both, partial credit is given.

Sample unit 3: TICKETS

In the unit TICKETS, students are invited to imagine that they have just arrived at a train station that has an automated ticketing machine. The context for the items in these units is classified as social and technological.

TICKETS: Stimulus information

TICKETS

A train station has an automated ticketing machine. You use the touch screen on the right to buy a ticket. You must make three choices.
- Choose the train network you want (subway or country).
- Choose the type of fare (full or concession).
- Choose a daily ticket or a ticket for a specified number of trips. Daily tickets give you unlimited travel on the day of purchase. If you buy a ticket with a specified number of trips, you can use the trips on different days.

The BUY button appears when you have made these three choices. There is a CANCEL button that can be used at any time BEFORE you press the BUY button.

At the machine, students can buy subway or country train tickets, with full or concession fares; they can choose daily tickets or a ticket for a specified number of trips. All items in this unit present an interactive problem situation: students are required to engage with the unfamiliar machine and to use the machine to satisfy their needs.

TICKETS: Item 1

Question 1: TICKETS CP038Q02

Buy a full fare, country train ticket with two individual trips.

Once you have pressed BUY, you cannot return to the question.
In the first item in the unit, students are invited to buy a full fare, country train ticket with two individual trips. This item measures the process of planning and executing. Students first have to select the network (“country trains”), then the fare type (“full fare”), then choose between a daily ticket and one for multiple individual trips, and finally indicate the number of trips (two). The solution requires multiple steps, and instructions are not given in the same order as they need to be applied. This is a relatively linear problem, compared to the following ones, but it is the first encounter with this new machine, which increases its level of difficulty relative to the following ones.

**TICKETS: Item 2**

### Question 2: TICKETS CP038Q01

You plan to take four trips around the city on the subway today. You are a student, so you can use concession fares.

Use the ticketing machine to find the cheapest ticket and press BUY.

Once you have pressed BUY, you cannot return to the question.

In the second item in the unit, students are asked to find and buy the cheapest ticket that allows them to take four trips around the city on the subway, within a single day. As students, they can use concession fares. This item is classified as exploring and understanding because this is the most crucial problem-solving process involved. Indeed, to accomplish the task, students must use a targeted exploration strategy, first generating at least the two most obvious possible alternatives (a daily subway tickets with concession fares, or an individual concession fare ticket with four trips), then verifying which of these is the cheapest ticket. If students visit both screens before buying the cheapest ticket (which happens to be the individual ticket with four trips) they are given full credit. Students who buy one of the two tickets without comparing the prices for the two only earn partial credit. Solving this problem involves multiple steps.

**TICKETS: Item 3**

### Question 3: TICKETS CP038Q03

You want to buy a ticket with two individual trips for the city subway. You are a student, so you can use concession fares.

Use the ticketing machine to purchase the best ticket available.

In the third item, students are asked to buy a ticket for two individual trips on the subway. They are told that they are eligible for concession fares. The third item in the unit is classified as monitoring and reflecting, since it requires them to modify their initial plan (to buy concession-fare tickets for the subway). When concession fares are selected, the machine says that “there are no tickets of this type available”. In this task, students must realise that it is not possible to carry through their initial plan, and so must adjust this plan by buying a full fare ticket for the subway instead.
Sample unit 4: TRAFFIC

TRAFFIC: Stimulus information

TRAFFIC

Here is a map of a system of roads that links the suburbs within a city. The map shows the travel time in minutes at 7:00 am on each section of road. You can add a road to your route by clicking on it. Clicking on a road highlights the road and adds the time to the Total Time box. You can remove a road from your route by clicking on it again. You can use the RESET button to remove all roads from your route.

In the unit TRAFFIC, students are given a map of a road network with travel times indicated. While this is a unit with static items, because all the information about travel times is provided at the outset, it still exploits the advantages of computer delivery. Students can click on the map to highlight a route, with a calculator in the bottom left corner adding up travel times for the selected route. The context for the items in this unit is classified as social and non-technological.

TRAFFIC: Item 1

Question 1: TRAFFIC CP007Q01
Pepe is at Sakharov and wants to travel to Emerald. He wants to complete his trip as quickly as possible. What is the shortest time for his trip?

- 20 minutes
- 21 minutes
- 24 minutes
- 28 minutes

In the first item in the unit, a planning and executing item, students are asked about the shortest time to travel from “Sakharov” to “Emerald”, two relatively close points shown on the map. Four response options are provided.

TRAFFIC: Item 2

The second item in the unit TRAFFIC is a similar planning and executing item. It asks students to find the quickest route between “Diamond” and “Einstein”, two distant points on the map. This time, students must provide their answer by highlighting this route. Students can use the indication that the quickest route takes 31 minutes to avoid generating all possible alternatives systematically; instead, they can explore the network in a targeted way to find the route that takes 31 minutes.
Assessing Problem-solving skills in Pisa 2012

Traffic: Item 2

**Question 2: TRAFFIC CP007Q02**
Maria wants to travel from Diamond to Einstein. The quickest route takes 31 minutes. Highlight this route.

Traffic: Item 3

**Question 3: TRAFFIC CP007Q03**
Julio lives in Silver, Maria lives in Lincoln and Don lives in Nobel. They want to meet in a suburb on the map. No-one wants to travel for more than 15 minutes. Where could they meet?

In the third item, students have to use a drop-down menu to select the meeting point that satisfies a condition on travel times for all three participants in a meeting. The demand in this third item is classified as a monitoring and reflecting task, because students have to evaluate possible solutions against a given condition.

Sample unit 5: ROBOT CLEANER

**Sample unit 5: ROBOT CLEANER**

The animation shows the movement of a new robotic vacuum cleaner. It is being tested.

Click the START button to see what the vacuum cleaner does when it meets different types of objects.

You can use the RESET button to place the vacuum cleaner back in its starting position at any time.
The unit ROBOT CLEANER presents students with an animation showing the behaviour of a robot cleaner in a room. The robotic vacuum cleaner moves forward until it meets an obstacle, then behaves according to a few, deterministic rules, depending on the kind of obstacle. Students can run the animation as many times as they wish to observe this behaviour. Despite the animated task prompt, the problem situations in this unit are static, because the student cannot intervene to change the behaviour of the vacuum cleaner or aspects of the environment. The context for the items in these units is classified as social and non-technological.

ROBOT CLEANER: Item 1

Question 1: ROBOT CLEANER CP002Q08
What does the vacuum cleaner do when it meets a red block?
- It immediately moves to another red block.
- It turns and moves to the nearest yellow block.
- It turns a quarter circle (90 degrees) and moves forward until it meets something else.
- It turns a half circle (180 degrees) and moves forward until it meets something else.

In the first item, students must understand the behaviour of the vacuum cleaner when it meets a red block. The item is classified as exploring and understanding. To show their understanding, they are invited to select, among a list of four options and based on observation, the description that corresponds to the behaviour of the robot cleaner in this situation: “It turns a quarter circle (90 degrees) and moves forward until it meets something else.”

ROBOT CLEANER: Item 2

Question 2: ROBOT CLEANER CP002Q07
At the beginning of the animation, the vacuum cleaner is facing the left wall. By the end of the animation it has pushed two yellow blocks. If, instead of facing the left wall at the beginning of the animation, the vacuum cleaner was facing the right wall, how many yellow blocks would it have pushed by the end of the animation?
- 0
- 1
- 2
- 3

In the second item in this unit, students must predict the behaviour of the vacuum cleaner using spatial reasoning. How many obstacles would the vacuum cleaner encounter if it started in a different position? This item is also an exploring and understanding item, because the correct prediction of the robot’s behaviour requires at least a partial understanding of the rules and careful observation of the animation to grasp the information needed. It is made easier if the student notes that the new starting position corresponds to an intermediate state of the robot’s trajectory in the animation. Response options are provided.

ROBOT CLEANER: Item 3

The final item in this unit is classified as representing and formulating, and asks students to describe the behaviour of the robot cleaner when it meets a yellow block. In contrast to the first task, students must formulate the answer themselves.
by entering it in a text box. This item requires expert scoring for credit. Full-credit answers are those that describe both of the rules that govern the robot’s behaviour (e.g. “it pushes the yellow block as far as it can and then turns around”). Partial credit was available for answers that only partially describe the behaviour, e.g. by listing only one of the two rules. Only a small percentage of students across participating countries obtained full credit for this item.
Notes

1. An assessment of collaborative problem-solving skills, which will be included in PISA 2015, will enrich the understanding of young people’s ability to solve problems.

2. Ramalingam, McCrae and Philpot (forthcoming) trace the history of how the PISA assessment of problem solving was developed and discuss its relationship with the psychological literature on problem solving and how it is measured.

References


http://dx.doi.org/10.1787/9789264006430-en  


http://dx.doi.org/10.1787/9789264180789-en  
