PISA 2012 FIELD TRIAL PROBLEM SOLVING FRAMEWORK

DRAFT SUBJECT TO POSSIBLE REVISION AFTER THE FIELD TRIAL

Doc: ProbSolvFrmwrk_FT2012

30 September 2010

Consortium:

Australian Council for Educational Research (ACER, Australia)

cApStAn Linguistic Quality Control (Belgium)

Deutsches Institut für Interup to nationale Pädagogische Forschung (DIPF, Germany)

Educational Testing Services (ETS, USA)

Institutt for Lærerutdanning og Skoleutvikling (ILS, Norway)

Leibniz - Institute for Science Education (IPN, Germany)

National Institute for Educational Policy Research (NIER, Japan)

The Tao Initiative: CRP - Henri Tudor and Université de Luxembourg - EMACS (Luxembourg)

Unité d'analyse des systèmes et des pratiques d'enseignement (aSPe, Belgium)

Westat (USA)

Framework for PISA 2012 Problem Solving

30 September 2010

TABLE OF CONTENTS

Framework for PISA 2012 Problem Solving
30 September 2010

PREAMBLE

1. This document presents the recommended framework for assessment of problem solving in the PISA 2012 field trial. The framework will be fine-tuned for the main survey following consideration of the outcomes of the field trial, and sample items from the field trial will be included as additional examples. The following paragraphs describe the development of the framework to this stage.

2. A first draft of the Problem Solving Framework was considered at the first Problem Solving Expert Group (PEG) meeting held in Melbourne from 10–12 February 2010. A second draft was prepared by Consortium staff immediately following the PEG meeting for presentation to the National Project Managers' meeting held in Hong Kong from 1–5 March. Feedback received at the Hong Kong NPM meeting, together with extensive written feedback from PEG members, was used to prepare a third draft intended for consideration at the PISA Governing Board meeting which was to be held in Copenhagen from 19–21 April 2010.

3. With the grounding of European flights due to volcanic ash cloud, the Copenhagen meeting was cancelled and replaced with a written consultation on issues that were to be settled at the meeting. As a consequence, comments from PGB members on the Problem Solving Framework were not available when the fourth draft was prepared for consideration at the PEG meeting held in Boston from 21–23 June 2010. However, the fourth draft did incorporate feedback on the third draft from PEG members and test developers. Only a small number of changes were made to the previous draft but some comments relating to unresolved matters, and to new issues raised by PEG members, were included for discussion in Boston.

4. Feedback from PGB members on the fourth draft was considered by the PEG at its Boston meeting, along with recommendations from the PISA Strategic Development Group (SDG) presented by Eugene Owen. Following the PEG meeting, Consortium staff prepared a fifth draft reflecting decisions made in response to the PGB and SDG feedback and the comments contained in the fourth draft. This fifth draft was circulated to PEG members for review and their feedback was incorporated in the version submitted to the OECD at the start of August. The present version incorporates changes agreed to at the PEG meeting held in Budapest from 27–29 September.

CHAPTER 1: INTRODUCTION

1. Problem solving competency is a central objective within the educational programmes of many countries. The acquisition of increased levels of problem solving competency provides a basis for future learning, for effective participation in society and for conducting personal activities. Students need to be able to apply what they have learned to new situations. The study of individuals' problem solving strengths provides a window on their capabilities to employ basic thinking and other general cognitive approaches to confronting challenges in life (Lesh & Zawojewski, 2007).

2. Problem Solving was an additional assessment domain in PISA 2003. Some key findings of the survey were as follows (OECD, 2004):

   - In some countries 70% of students could solve relatively complex problems, while in others less than 5% could do so.
   - In most countries, more than 10% of students were unable to solve basic problems.
   - On average in OECD countries, half of the students were unable to solve problems that are more difficult than basic problems.
   - Patterns of within-country variation in students' problem solving proficiency differed considerably across countries.
   - Patterns of within-country differences between problem-solving proficiency and domain-related proficiencies (mathematics, reading, science), differed considerably across countries.

3. Since the 2003 problem solving assessment framework (OECD, 2003a) was developed, considerable research has been carried out in the areas of complex problem-solving, transfer, computer based assessment of problem solving, and large-scale assessment of problem solving competency (*e.g.* Blech & Funke, 2005; Funke & Frensch, 2007; Greiff & Funke, 2008; Klieme, 2004; Klieme, Leutner, & Wirth, 2005; Leutner, Klieme, Meyer, & Wirth, 2004; Mayer, 2002; Mayer & Wittrock, 2006; O'Neil, 2002; Osman, 2010; Reeff, Zabal & Blech, 2006; Wirth & Klieme, 2004). This research has led to advances in understanding and measuring individuals' problem solving capabilities.

4. In addition, advances in software development tools and the use of networked computers have made possible greater efficiency and effectiveness of assessment, including the capability to administer dynamic and interactive problems, engage students' interest more fully and capture more information about the course of the problem-solving process. On this last point, computer delivery of assessment tasks makes it possible to record data about such things as the type, frequency, length and sequence of actions performed by students in responding to items.

5. It is appropriate, therefore, to once again make problem solving an assessment domain in PISA, but in doing so to devise a new framework and implement additional assessment methodologies that allow for the real-time capture of students' capabilities. In particular, the PISA 2012 assessment of problem solving will be computer-based and interactivity of the student with the problem will be a central component of the information gathered.

6. PISA 2012 problem solving is an assessment of *individual* problem solving competency. Collaborative problem solving skills – the skills required to solve problems as a member of a group – are essential for successful employment, where the individual is often a member of a team of diverse specialists working in separate locations. However, significant measurement challenges still stand in the way of collaborative tasks becoming a feature of large-scale international surveys such as PISA (Reeff, Zabal & Blech, 2006). Foremost among these challenges are how to assign credit to individual group members if this is required, how to account for differences across groups that may bias individual performance, and how to account for cultural differences in group dynamics.

7. A consistent research finding is that problem solving is dependent on domain-specific knowledge and strategies (*e.g.* Mayer, 1992; Funke & Frensch, 2007)[1]. The PISA 2012 assessment will avoid the need for expert prior knowledge as much as possible in order to focus on measuring the cognitive processes fundamental to problem solving. This also distinguishes the assessment from problem solving tasks in the core PISA literacy domains of reading, mathematics and science, which call on expert knowledge in these areas.

8. Another conclusion that can be drawn from the research overview is that authentic, relatively complex problems, particularly those that require direct interaction by the solver to uncover and discover relevant information, should be a

[1] An overview of scientific research into individual problem solving is included as Annex A. This review need not necessarily be included in the final framework publication.

central feature of the PISA 2012 problem solving assessment. Examples are the problems commonly faced when using unfamiliar everyday devices such as remote controls, personal digital devices (*e.g.* mobile phones), home appliances and vending machines. Other examples arise in situations such as physical conditioning, feeding animals, growing plants and social interactions. Problem-solving skills are necessary to achieve more than a basic level of skill in dealing with such situations and there is evidence that skills additional to those involved in traditional reasoning-based problem solving are required (*e.g.* Klieme, 2004). This will be the first time, made possible by computer delivery of the assessment, that such "interactive problems" have been included in a large-scale international survey.

9. Problem solving competency can be developed by high quality education. 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 and develops the cognitive processes that underpin problem solving. It prepares students to reason effectively in unfamiliar situations, and to fill gaps in their knowledge by observation, exploration and interaction with unknown systems. The PISA 2012 computer-based assessment of problem solving aims to examine how students are prepared to meet unknown future challenges for which direct teaching of today's knowledge is not sufficient.

Problem Solving in PIAAC

10. The OECD's Programme for the International Assessment of Adult Competencies (PIAAC), is an assessment of reading component skills, literacy, numeracy, and problem solving in technology-rich environments. It is a face-to-face sample household survey of people aged 16–65 years and will be conducted for the first time in 2012.

11. PIAAC's assessment of "problem solving in technology rich environments" differs from the PISA 2012 assessment of problem solving in two important aspects[2]. Firstly, it is primarily concerned with "information-rich" problems. Examples include needing to locate and evaluate information on the Web or on social

[2] PIAAC defines *problem solving in technology rich environments* as follows: "Problem solving in technology rich environments involves using digital technology, communication tools and networks to acquire and evaluate information, communicate with others and perform practical tasks." (OECD, March 2009, p.7)

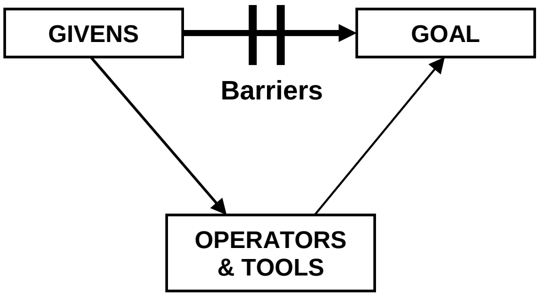
networking sites, navigating through unfamiliar web pages and making decisions about what information is relevant and irrelevant for a task.

12. A second major difference is that problem solutions require the use of one or more computer software applications (file management, web browser, email, and spreadsheet). In PISA, ICT is integral to the assessment of problem solving but it is not integral to its definition of problem solving. Only foundational ICT skills (based on use of a keyboard and mouse) are necessary to undertake the PISA assessment of problem solving. Software tools are common and powerful aids for information-rich problem solving and a high level of ICT literacy is essential in this digital age. However, PISA's focus is assessment of the fundamental cognitive processes that are essential for successful problem solving with or without ICT assistance.

13. The aim of the PISA 2012 problem solving assessment is to assess individual *problem solving competency*. Before defining what is meant by the term “problem solving competency” in this context, it is important to clarify what is meant by the terms “problem” and “problem solving” by researchers in the field.

14. A problem exists when a person has a goal but doesn’t know how to achieve it (Duncker, 1945). This definition is enlarged upon in *Figure 1*. The *given state* (givens) is the knowledge the person has about the problem at the outset and the *operators* are the admissible actions that can be performed to achieve the desired *goal state* (outcomes) with the assistance of the available *tools*. *Barriers* that must be overcome (*e.g.* lack of knowledge or obvious strategies) stand in the way of achieving the goal. Overcoming the barriers may involve not only cognition, but motivational and affective means (Funke, 2010).

Figure 1. **Problem Situation** (after Funke & Frensch, 1995)



15. As an example, consider the simple problem of finding the route between two towns that is likely to be the quickest, given a road map with estimated travel times marked and a calculator. The given state is the given information – the map with no route marked – and the goal state is the desired answer – the quickest route. The allowable actions (operators) are selecting a possible route, calculating its total time and comparing it with the times for other routes. A tool (calculator) is available for assistance in adding times.

16. Consistent with this understanding of what is meant by a problem, Mayer (1990) defines *problem solving* as cognitive processing directed at transforming a given situation into a goal situation when no obvious method of solution is available. This definition is widely accepted in the problem solving community (*e.g.* see Klieme, 2004; Mayer & Wittrock, 2006; Reeff, Zabal & Blech, 2006).

17. The PISA 2012 definition of **problem solving competency** is grounded in these generally-accepted meanings of "problem" and "problem solving". It is as follows:

- *Problem solving competency is 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.*

18. Not surprisingly, the first sentence of the definition is almost identical to the first part of the definition used for the PISA 2003 assessment of problem solving[3]. However, whereas the 2003 definition had only a cognitive dimension, with its latter part highlighting the cross-curricular nature of the assessment, an affective component is introduced in the 2012 definition in keeping with the definition of competency as recognised by the OECD (OECD, 2003a).

19. What distinguishes the 2012 assessment of problem solving from the 2003 assessment is not so much the definition of problem solving competency, but the mode of delivery (computer-based) of the 2012 assessment and the inclusion of problems which cannot be solved without the solver *interacting* with the problem situation.

20. In the following paragraphs, each part of the PISA 2012 definition of problem solving competency is considered in turn to help clarify its meaning in relation to the assessment.

Problem solving competency…

21. A *competency* involves far more than the basic reproduction of accumulated knowledge. It involves a mobilisation of cognitive and practical skills, creative abilities and other psychosocial resources such as attitudes, motivation and values (OECD, 2003b). The PISA 2012 assessment of problem solving competency will not test simple reproduction of domain-based knowledge; rather it will focus on the cognitive skills required to solve unfamiliar problems encountered in life[4] and lying outside traditional curricular domains.

[3] "Problem solving is an individual's capacity to use cognitive processes to confront and resolve real, cross-disciplinary situations where the solution path is not immediately obvious and where the literacy domains or curricula areas that might be applicable are not within a single domain of mathematics, science or reading" (OECD, 2003a, p. 156).
[4] Including educational and professional encounters.

22. Prior knowledge is important in solving problems. However, problem solving competency involves the ability to acquire and use new knowledge, or to use old knowledge in a new way, to solve novel problems (*i.e.* problems that are not routine).

...is an individual's capacity to engage in cognitive processing...

23. Problem solving occurs internally in an individual's cognitive system and can only be inferred indirectly by the person's actions and products. It involves representing and manipulating various types of knowledge in the problem solver's cognitive system (Mayer & Wittrock, 2006). Students' responses to assessment items – their exploration strategies, the representations they employ in modelling the problem, numerical and non-numerical answers, or extended explanations of how a problem was solved – will be used to make inferences about the cognitive processes they employed.

24. Creative (divergent) thinking and critical thinking are important components of problem solving competency (Mayer, 1992). Creative thinking is a cognitive activity that results in finding solutions to a novel problem. Critical thinking accompanies creative thinking and is employed to evaluate possible solutions. The assessment will target both components.

...to understand and resolve problem situations...

25. To what degree can individuals meet the challenges of a problem situation and move towards resolving it? In addition to explicit responses to items, the assessment aims to measure individuals' progress in solving a problem, including the strategies they employ. Where appropriate these strategies can be tracked by means of behavioural data captured by the computer-delivery system: the type, frequency, length and sequence of interactions made with the system can be captured and used in scoring or in subsequent analyses of student performance.

26. 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 and to plan and carry out a solution, along with monitoring and evaluating progress throughout the activity.

27. Often in real-world problems there may not be a unique or exact solution. In addition, the problem situation may change during the solving process, possibly due to interaction with the problem solver or as a result of its own dynamic

nature. These complexities will be addressed when setting assessment tasks and a balance struck between authenticity of a situation and practicality of assessment.

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

28. The means of finding a solution path should not be immediately obvious to the problem solver. There will be barriers of various sorts in the way, or missing information. The assessment is concerned with *non-routine* problems, not routine ones (*i.e.* problems for which a previously learned solution procedure is clearly applicable): the problem solver must actively explore and understand the problem and either devise a *new* strategy or apply a strategy learnt *in a different context* to work towards a solution.

29. The status of a problem – whether it is routine or not – depends on the solver's familiarity with the problem. A "problem" for one person may have an obvious solution to another person who is experienced and practised with such problems. Accordingly, care will be taken to set problems that should be non-routine for the great majority of 15-year-olds.

30. It is not necessarily the case that the context or goals will themselves be unfamiliar to the solver; what is important is that the particular problems are novel or the ways of achieving the goals are not immediately obvious. The problem solver might need to explore or interact with the problem situation before attempting to solve the problem. Direct interaction is made feasible by the use of computer-delivered assessment in PISA 2012.

…It includes the willingness to engage with such situations…

31. Problem solving is personal and directed, that is, the problem solver's processing is guided by their personal goals (Mayer & Wittrock, 2006). The problem solver's individual knowledge and skills help determine the difficulty or ease with which obstacles to solutions can be overcome. However, the operation of such knowledge and skill is affected by motivational and affective factors such as beliefs (*e.g.* self-confidence) and feelings about one's interest and ability to solve the problem (Mayer, 1998).

32. In addition, the context of a problem (whether it is familiar and understood), the external resources available to the solver (such as access to tools), and the environment in which the solver operates (*e.g.* an examination setting) will affect the way a person approaches and engages with the problem.

33. Motivation and affective factors will not be measured in the problem solving assessment but some may be addressed generally, or with particular reference to mathematics, in the student questionnaire.

…in order to achieve one's potential as a constructive and reflective citizen.

34. Competency is an important factor in the ways that individuals help to shape the world, not just to cope with it: "…key competencies can benefit both individuals and societies" (Rychen & Salganik, 2003). They should "manage their lives in meaningful and responsible ways by exercising control over their living and working conditions" (*ibid*). Individuals need to be proficient problem solvers to achieve their potential as constructive, concerned and reflective citizens.

Scope of the Assessment

35. The PISA 2012 assessment of problem solving will not include problems that require expert prior knowledge for their solution. In particular, problems that could reasonably be included in an assessment of one of the three core PISA domains will not be included. Assessment tasks will centre on everyday situations, with a wide range of contexts employed as a means of controlling for prior knowledge in general.

36. Mobilisation of prior knowledge is not sufficient to solve novel problems in many everyday situations. Gaps in knowledge must be filled by observation and exploration of the problem situation. This often involves interaction with a new system to discover rules that in turn must be applied to solve the problem. Instead of a straightforward application of previously mastered knowledge, existing knowledge needs to be reorganised and combined with new knowledge using a range of reasoning skills.

CHAPTER 3: ORGANISING THE DOMAIN

37. How the domain is represented and organised determines the assessment design and, ultimately, the evidence about student proficiencies that can be collected and reported. Many elements are part of the construct, not all of which can be taken into account and varied in an assessment such as PISA. It is necessary to select the most important elements that can be varied to ensure construction of an assessment that contains items which have an appropriate range of difficulty and provide a broad coverage of the domain.

38. The domain elements of key importance for the PISA 2012 assessment of problem solving are as follows:

- The *problem context*: whether it involves a technological device or not, and whether the focus of the problem is personal or social.
- The *nature of the problem situation*: whether it is interactive or static.
- The *problem solving processes*: the cognitive processes involved in solving a problem.

39. Tasks will be constructed to measure how well students perform when the various cognitive processes involved in problem solving are exercised within the two different types of problem situations across a range of contexts. Each of these key domain elements is discussed and illustrated in the following sections. The characteristics of the tasks themselves are discussed in Chapter 4.

a. Problem Context

40. An individual's familiarity and understanding of the problem context will affect how difficult the problem is to solve for that person. Two dimensions have been identified to ensure that assessment tasks sample across a range of contexts that are authentic and of interest to 15-year-olds: the setting (technology or not) and the focus (personal or social).

41. Problems set in a *technology context* have the functionality of a technological device as their basis. Examples include mobile phones, remote controls for appliances and ticket vending machines. Knowledge of the inner workings of these devices will not be required: typically, problem solvers will be led to explore and understand the functionality of a device, as preparation for controlling the device or for troubleshooting its malfunctioning. Situations that give rise to other types of

problems, such as route planning, task scheduling and decision making, have *non-technology* contexts.

42. *Personal* contexts include those relating primarily to the self, family and peer groups. *Social* contexts relate to situations encountered in the community or society in general. As illustrations, the context of an item about setting the time on a digital watch would be classified as Technology and Personal, whereas the context of an item requiring the construction of a basketball team roster would be classified as Non-technology and Social.

b. Nature of Problem Situation

43. How a problem is presented has important consequences for how it can be solved. Of crucial importance is whether the information about the problem disclosed to the problem solver at the outset is complete. This is the case for the quickest-route problem discussed earlier (see paragraph 15). We refer to such problem situations as being *static*.

44. By contrast, problem situations may be *interactive*, meaning that exploration of the situation to uncover additional relevant information is possible[5]. Real-time navigation using a GPS system, where traffic congestion is reported automatically or by query, presents such a situation.

45. Interactive problem situations can be simulated in a test setting by means of a computer. Including interactive problem situations in the computer-based PISA 2012 problem solving assessment allows a wide range of more authentic, real-life scenarios to be presented than would otherwise be possible using pencil-and-paper tests. Problems where the student explores and controls a simulated environment will be a distinctive feature of the assessment.

46. A selection of static problem situations also will be included in the assessment. The assessment of such problems has traditionally taken place using pencil-and-paper tests. However, their computer-based assessment has many advantages including the capability of presenting a broader range of scenarios, involving multimedia elements such as animation; the availability of on-line tools; and, the use of a wide range of response formats that can be automatically coded.

[5] The term "intransparent" is sometimes used to describe problems when complete information about the problem situation is not available at the outset (see Funke & Frensch, 1995).

47. Some studies suggest that knowledge acquisition in exploring a problem in an interactive environment, and how that knowledge is applied, are competencies distinct from the typical skills used in solving static problems (see Klieme, 2004; Wirth & Klieme, 2004; Leutner & Wirth, 2005). Including a mixture of interactive and static problems in the PISA 2012 assessment will provide a broader measure of problem-solving competency than has been possible with pencil-and-paper instruments.

Interactive Problem Situations

48. Interactive problem situations often arise when encountering real-world artefacts such as ticket vending machines, air-conditioning systems or mobile telephones for the first time, especially if the instructions for use of such devices are not clear or not available. Understanding how to control such devices is a problem faced universally in everyday life. In these situations it is often the case that some relevant information is not apparent at the outset. For example, the effect of applying an operation (say, pushing a button on a remote control) may not be known and cannot be deduced, but rather must be inferred by interacting with the scenario through actually performing the operation (pushing the button) and forming a hypothesis about its function based on the outcome. In general, some exploration or experimentation must be done to acquire the knowledge necessary to control the device. Another common scenario is when a person must troubleshoot a fault or malfunction in a device. Here a certain amount of experimentation must take place to collect data on the circumstances under which the machinery fails.

49. An interactive problem situation can be *dynamic*, meaning that its state might change of its own accord due to influences beyond the problem solver's control (*i.e.* without any intervention by the problem solver)[6]. For example, in the case of a ticket-vending machine, if during a transaction no buttons are pressed for 20 seconds the machine might reset. Such autonomous behaviour of a system must be observed and understood so it can be taken into account in attaining the desired goal (purchasing a ticket).

[6] The term "dynamic" is used by some researchers to describe any simulated physical system that a problem solver can interact with and receive feedback. In such cases, problem situations that change autonomously are sometimes termed "eigendynamic" (*e.g.* see Blech & Funke, 2005).

Static Problem Situations

50. Static problem situations can give rise to *well-defined* or *ill-defined* problems. In a *well-defined problem*, such as the quickest-route problem (see paragraph 15), the given state, goal state and allowable operators are clearly specified (Mayer & Wittrock, 2006). The problem situation is not dynamic (i.e. does not change of its own accord during the course of solving the problem), all relevant information is disclosed at the outset and there is a single goal.

51. Other examples of well-defined problems are traditional *logic puzzles* such as the Tower of Hanoi and the water jars problems (see, for example, Robertson, 2001); *decision-making problems*, where the solver is required to understand a situation involving a number of well-defined alternatives and constraints so as to make a decision that satisfies the constraints (*e.g.* choosing the right pain killer, given sufficient details about the patient, the complaint and the available pain killers); and, *scheduling problems* for projects such as building a house or producing computer software, where a list of tasks with durations and dependencies between tasks is given.

Ill-defined Problems

52. Mayer and Wittrock (2006) point out that "educational materials often emphasise well-defined problems, although most real problems are ill-defined [*i.e.* not well-defined]". These problems often involve multiple goals which are in conflict so that progress toward one may detract from progress toward the other(s). Elaboration and weighing of priorities is required for the problem solver to achieve a balance between the goals (Blech & Funke, 2010). An example is finding the "best" route between two places – should this be the shortest route?, the likely quickest route?, the most straightforward route?, the route with minimum variation in time?, *etc*. A more complex example is designing a car where high efficiency, low cost, high safety and low effect on the environment are desired.

c. Problem Solving Processes

53. Different authors conceive of the cognitive processes involved in solving a problem in different ways, but there is a great deal of commonality in their views. The processes identified below are derived from the work on problem solving and reasoning of cognitive psychologists (*e.g.*, Baxter & Glaser, 1997; Bransford *et al*, 1999; Mayer & Wittrock, 1996; Mayer & Wittrock, 2006; Vosniadou & Ortony, 1989), as well as by the seminal work of Polya (1945). Additionally, recent work

on complex and dynamic problem solving (Blech & Funke, 2005, 2010; Funke & Frensch, 2007; Greiff & Funke 2008; Klieme, 2004; Osman, 2010; Reeff, Zabal & Blech, 2006; Wirth & Klieme, 2004) has been taken into account.

54. No assumption is made that the processes involved in solving a particular problem are sequential or that all of the processes listed are necessary in the solution of a particular problem. As individuals confront, structure, represent and solve authentic problems representing emerging life demands, they may move to a solution in a way that transcends the boundaries of a linear, step-by-step model. Most of the information about the functioning of the human cognitive system now supports the view that it is capable of parallel information processing (Lesh & Zawojewski, 2007).

55. For the purposes of the PISA 2012 problem solving assessment, the processes involved in problem solving are taken to be:

- Exploring and understanding
- Representing and formulating
- Planning and executing
- Monitoring and reflecting

56. **Exploring and understanding.** The objective here is to build mental representations of each of the pieces of information presented in the problem. This involves:

- *exploring* the problem situation: observing it; interacting with it; searching for information; finding limitations or obstacles; and
- *understanding* given information and information discovered while interacting with the problem situation; demonstrating understanding of relevant concepts.

57. **Representing and formulating.** The objective here is to build a coherent mental representation of the problem situation (*i.e.* a situation model or a problem model). To do this, relevant information must be selected, mentally organised, and integrated with relevant prior knowledge. This may involve:

- *representing* the problem by constructing tabular, graphical, symbolic or verbal representations, and shifting between representational formats; and

- *formulating* hypotheses by identifying the relevant factors in the problem and their interrelationships; organising and critically evaluating information.

58. **Planning and executing.** This includes:

- *planning*, which consists of goal setting, including clarifying the overall goal, and setting sub-goals, where necessary; and devising a plan or strategy to reach the goal state, including the steps to be undertaken; and

- *executing*, which consists of carrying out a plan.

59. **Monitoring and reflecting** includes:

- *monitoring* progress towards the goal at each stage, including checking intermediate and final results, detecting unexpected events, and taking remedial action when required; and

- *reflecting* on solutions from different perspectives; critically evaluating assumptions and alternative solutions; and looking for additional information or clarification.

Reasoning Skills

60. Each of the problem solving processes draws upon one or more *reasoning skills*. In understanding a problem situation, the problem solver may need to distinguish between facts and opinion; in formulating a solution, the problem solver may need to identify relationships between variables; in selecting a strategy, the problem solver may need to consider cause and effect; and, in communicating the results, the problem solver may need to organise information in a logical manner. The reasoning skills associated with these processes are embedded within problem solving. They are important in the PISA context since they can be taught and modelled in classroom instruction (*e.g.* Adey *et* al, 2007; Klauer & Phye, 2008).

61. Examples of reasoning skills employed in problem solving include deductive, inductive, quantitative, correlational, analogical, combinatorial, and multidimensional reasoning. These reasoning skills are not mutually exclusive and often in practice problem-solvers move from one to another in gathering evidence and testing potential solution paths before settling into the major use of one

method over others in finding the solution to a given problem. Reasoning skills will be broadly sampled when devising assessment items as they are likely to influence the difficulty of items.

CHAPTER 4: ASSESSING PROBLEM SOLVING COMPETENCY

a. Structure of the Assessment

62. The duration of the main survey assessment will be 40 minutes. There will be a total of 80 minutes of material organised into four 20-minute clusters, with each student doing two clusters according to a balanced rotation design. About twice this amount of material will be included in the field trial. It is estimated that 20–25 items will be included in the main survey, with about one-third of them being partial credit items. Timing information automatically captured during the field trial will be used to determine the actual number of items that can be included.

63. As is normal for PISA assessments, items will be grouped in units based around a common stimulus that will describe the problem situation. To minimise the level of reading literacy required, stimulus material (and task statements) will be as clear, simple and brief as possible. Animations, photographs or diagrams will be used to avoid lengthy passages of text. Numeracy demands also will be kept to a minimum with, for example, running totals provided where appropriate.

64. The assessment will include a broad sample of items covering a range of difficulty that will enable the strengths and weaknesses of populations and key subgroups to be determined with respect to the cognitive processes involved in problem solving.

Functionality Provided by Computer Delivery

65. A principal benefit of measuring problem solving competency by computer is the opportunity to collect and analyse data that relate to processes and strategies, in addition to capturing and scoring intermediate and final results. This is likely to be a major contribution of the PISA 2012 assessment of problem solving. With appropriate item authoring, data such as the type, frequency, length and sequence of actions performed by students can be captured for this purpose.

66. Another benefit is that the time students spend on any particular item can be monitored and restricted. This facility may be used where it is considered appropriate, such as in limiting the time students are allowed to explore a complex problem situation (or, to put it another way, for limiting the time students may *waste* in this manner).

67. Only foundational ICT skills will be assumed in the assessment, such as keyboard use, manipulating a pointer (*e.g.* a mouse), clicking buttons, drag-and-drop, scrolling, and use of pull-down menus and hyperlinks. Care will be taken to ensure

that interference with the measurement of problem solving competency by ICT demand and presentation is kept to a minimum.

68. Both units and items within units will be delivered in a fixed order, or "lockstep" fashion. The lockstep procedure means that students are not able to return to an item or unit once they have moved to the next one. Each time students click the Next button a dialog box will display a warning that they are about to move on to the next item and that it will not be possible to return to the previous item. At this point, students can either confirm they want to move on or cancel the action and return to the current item.

b. Task Characteristics and Difficulty

69. Generally, each item will focus on a single problem solving process as far as possible. Accordingly, for some items, demonstrating a recognition of the problem will be sufficient; in others, describing a method of solution will be enough; in many, the actual solution(s) will be required with effectiveness and efficiency of method being important characteristics; in yet others, the task will be to evaluate proposed solutions and decide on the most appropriate solution for the problem posed. Although executing is often emphasised in classroom instruction, the major difficulties for most problem solvers involve representing, planning and self-regulating (Mayer, 2003).

70. Some problems are inherently more complex than others (Funke & Frensch, 2007). Furthermore, increased complexity generally means greater difficulty. *Table 1* summarises task characteristics that will be varied in the assessment to ensure that the items cover an appropriate range of difficulty.

***Table 1*. Task Characteristics**

Characteristic	Effect on task difficulty
Amount of information	The more information that has to be considered, the more difficult the task is likely to be.
Representation of information	Unfamilar representations, and multiple representations (especially if the information has to be related), tend to increase difficulty.
Disclosure of information	The more relevant information that is not disclosed at the outset and therefore has to be discovered (*e.g.* effect of operations, autonomous behaviour, unanticipated obstacles), the more difficult the task is likely to be.
Internal complexity	Internal complexity of a problem situation increases as the number of variables and their dependencies increases.Tasks with a high level of internal complexity are likely to be harder than those with lower levels of internal complexity.
Constraints to be satisfied	Tasks in which there are few constaints on the values of variable are likely to be easier than those in which many constraints must be satisfied.
Distance to goal	The greater the number of steps needed to solve a problem, the more difficult it is likely to be.
Reasoning skills required	Problems that require the application of some types of reasoning skills (*e.g.* combinatorial reasoning) are likley to be harder than those that do not.

Response Formats and Coding

71. At least 75% of items will have a response format that can be automatically coded in its captured form. This will include simple and complex multiple-choice items that are answered by clicking option (radio) buttons, items that require shapes to be selected and dragged into position, items in which selections have to be made from pull-down menus to fill table cells, and items that require parts of diagrams to be highlighted.

72. Open constructed-response items will be used where necessary, for example where it is considered important to ask students to explain their method or justify their solution. For these items, students will enter their responses in text boxes.

73. Constructed responses will be collected automatically by the computer-delivery system and an Online Coding System will be developed to facilitate their coding by experts. This eliminates the need for separate data entry, minimises the need for data cleaning, and allows coding to take place “off site” if desired.

74. The coding scheme for an item will allow for *partial credit* if appropriate, such as when multiple correct answers are required for full credit or when a correct strategy is employed but is not executed properly. Designated behaviours (such as explorations strategies) that provide reliable evidence about problem solving competency over and above the fulfilment of task demands will be captured and contribute to scoring.

Interactive Problems

75. Interactive problems can be built on underlying mathematical models whose parameters can be varied systematically to achieve differing degrees of difficulty. There are two well-used paradigms: *linear difference equations* and *finite state machines*.

76. With problem situations modelled by linear difference equations (also referred to as linear structural equations)[7], the problem solver must manipulate one or more input variables (such as controls for a climate control system) and consider the effect this has on one or more output variables (such as temperature and humidity); the output variables may also influence themselves so that the system is dynamic. Example contexts include remote controls, thermostats, paint mixing and ecosystems.

77. A *finite state machine* is a system with a finite number of states, input signals and output signals (Buchner & Funke 1993)[8]. The system's next state (and output signal) is uniquely determined by its current state and the specific input signal. With problem situations modelled by finite state machines, the problem solver must supply input signals (usually in the form of a sequence of button presses) to determine the effect on the system's states in an effort to understand its underlying structure and move it towards a goal state. Many everyday devices and contexts are governed or constrained by the rules of a finite state machine structure. Examples include digital watches, mobile phones, microwave ovens, MP3-players, ticket vending machines and washing machines.

78. The typical task demands for such interactive problems are as follows (see Blech & Funke, 2005 and Greiff & Funke, 2008 for additional detail):

---

[7] See Greiff & Funke (2008) who use the term *MicroDYN* to describe these systems. An earlier implementation of such a system is known as *Dynamis* – see Blech & Funke (2005).

[8] Finite state machines for assessment purposes have been implemented under the name "*MicroFin*" – see http://www.psychologie.uni-heidelberg.de/ae/allg_en/forschun/probleml.html

- *Exploration*. Acquire knowledge of system structure either by active or directed exploration (interaction). *[Exploration strategies can be tracked and captured by the computer delivery system.]*

- *Identification*. Give or complete a representation of the mental model of the system that is formed during exploration. This may be in drawing or text form. *[The accuracy of the model helps in assessing acquired causal knowledge.]*

- *Control*. Practical application of acquired knowledge: transform a given state into a goal state and (for appropriate systems) maintain the goal state over time. A correct model of the system may be provided to minimise dependence on previous items. *[Transfer of acquired knowledge is assessed in this way.]*

- *Explanation*. Describe strategies used to reach a goal; explain how a system works; or suggest causes of a malfunction of a device.

79. Students may already have some idea of the relationships between system variables in problem situations because of their familiarity with similar, actual devices. Such prior knowledge will vary between individuals and so a *variety* of common, everyday problem contexts will be used to help overcome this effect across the assessment. In addition, a few more unusual but engaging game-like contexts will be included where the relationships must be inferred solely by manipulation and observation of system variables.

80. The difficulty of problems of these types is largely dependent on the internal complexity of the mathematical models underlying the situations. Problems of varying difficulty can be set by systematically varying this complexity which is determined by the number of variables involved and how they are connected. For example, a problem involving only a few variables can be very easy if it only involves direct effects between input and output variables, but can be made extremely difficult by the inclusion of multiple effects and side effects between output variables.

c. Distribution of Items

81. The recommended percentage distribution of score points according to the cognitive processes involved in problem solving is given in *Table 2*. Highest weighting is given to *Planning and executing* in recognition of the importance of being able to carry through a solution to a successful conclusion. Lower than average weight is given to *Monitoring and reflecting* because it is an integral part of

the other three processes and therefore also is assessed (indirectly) in items that target those processes.

Table 2. **Distribution of Score Points by Process**

Exploring & understanding	Representing & formulating	Planning & executing	Monitoring & reflecting	Total
20–25%	20–25%	35–45%	10–20%	*100%*

82. *Table 3* indicates the recommended distribution of items across the two other key domain elements, problem context and nature of the problem situation[9]. The clear emphasis on interactive problems over static problems (a ratio of about 2:1) reflects the intention that the assessment should concentrate on this important class of problems which, with the advantage of computer delivery, it is possible to include in a large-scale international survey for the first time.

Table 3. **Distribution of Items by Problem Nature and Context**

	Technology context	Non-technology context	
Static problem situation	5–10%	20–25%	25–35%
Interactive problem situation	40–45%	25–30%	65–75%
	45–55%	45–55%	*100%*

[9] Note that these two classifications will be assigned at the *unit* level and apply to all items in the unit.

CHAPTER 5: REPORTING PROBLEM SOLVING COMPETENCY[10]

83. It is expected that at least four levels of proficiency will be able to be identified and described to show how individuals' problem solving competency grows and develops, and to enable comparisons of student performance between and within participating countries and economies. There will not be enough items in the main survey to report on subscales.

84. Proficiency descriptions characterising typical student performance at each level will be developed by analysing the knowledge and skills required to answer the items at that level. It is expected that the following abilities will characterise high-performing students:

- Ability to plan and execute solutions that involve a relatively high number of steps, and to apply a wide range of reasoning skills.
- Ability to deal with situations involving many variables where there is high dependency between the variables or a large number of constraints on their values.
- Ability to understand and relate information presented in a variety of familiar or unfamilar representations.
- Ability to interact with problems to discover undisclosed information or handle unanticipated obstacles.

85. Low-performing students are expected to have the following characteristics:

- At best be able to plan and execute solutions that involve only a few steps and simple reasoning.
- Be unable to understand information presented in unfamiliar representations or to relate information between familiar representations.
- Be able to make little or no progress in solving a problem unless it involves only one or two variables with very limited dependency and no constraints.
- Only be able to discover undisclosed information if instructions are provided to direct exploration activity.

[10] Further detail will be added when field trial results are available.

CHAPTER 6: SAMPLE TASKS[11]

86. DIGITAL WATCH

A simulation of a digital watch is presented. The watch is controlled by four buttons, the functions of which are unknown to the student at the outset of the problem. The student is required to: (Q1) determine through guided exploration how the buttons work in TIME mode; (Q2) complete a diagram showing how to cycle through the various modes; and (Q3) use this knowledge to control the watch (set the time).

- *Problem Context*: Technology, Personal
- *Nature of Problem*: Interactive
- *Problem Solving Processes*: (Q1) Exploring and understanding; (Q2) Representing and formulating; (Q3) Planning and executing
- *Reasoning Skills*: Correlational, multidimensional

87. BASKETBALL

The rules for a basketball tournament relating to the way in which match time should be distributed between players are given. There are two more players than required (5) and each player must be on court for at least 25 of the 40 minutes playing time. Students are required to: (Q1) create a schedule for team members that satisfies the tournament rules; and (Q2) reflect on the rules by critiquing an existing schedule.

- *Problem Context*: Social, Non-technology
- *Nature of Problem:* Static
- *Problem Solving Process*: (Q1) Planning and executing; (Q2) Monitoring and reflecting
- *Reasoning Skills:* Combinatorial, deductive

[11] Field trial tasks not needed for the main survey will be added once the main survey selection is made, with some examples being distributed throughout the text and cross-referenced to illustrate the discussion.

REFERENCES

Adey, P., Csapo, B., Demetriou, A., Hautamäki, J. & Shayer, M. (2007). Can we be intelligent about intelligence? Why education needs the concept of plastic general ability. *Educational Research Review 2*, 75–97.

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J. & Wittrock, M. C. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.

Baxter, G. P. & Glaser, R. (1997). An approach to analysing the cognitive complexity of science performance assessments (Technical Report 452), National Center for Research on Evaluation, Standards and Student Testing (CRESST), Los Angeles, CA.

Blech, C. & Funke, J. (2005). Dynamis review: An overview about applications of the Dynamis approach in cognitive psychology. Bonn: Deutsches Institut fűr Erwachsenenbildung (available at http://www.die-bonn.de/esprid/dokumente/doc-2005/blech05_01.pdf).

Blech, C. & Funke, J. (2010). You cannot have your cake and eat it, too: How induced goal conflicts affect complex problem solving. *Open Psychology Journal 3*, 42–53.

Bransford, J. D., Brown, A. OL. & Cockling, R. R. (Eds.) (1999). *How People Learn: Brain, Mind, Experience, and School*. National Academy Press: Washington, DC.

Buchner, A. & Funke, J. (1993). Finite-state automata: Dynamic task environments in problem-solving research. *The Quarterly Journal of Experimental Psychology 46A* (1), 83–118.

Duncker, K. (1945). On problem solving. *Psychological Monographs 58* (3) (Whole No. 270).

Funke, J. (2001). Dynamic systems as tools for analysing human judgement. *Thinking and Reasoning, 2001, 7* (1), 69–89.

Funke, J. (2010). Complex problem solving: A case for complex cognition? *Cognitive Processing 11*, 133–142.

Funke, J. & Frensch, P. A. (2007). Complex problem solving: The European perspective – 10 years after. In D. H. Jonassen (Ed.), *Learning to Solve Complex Scientific Problems* (pp. 25–47). New York: Lawrence Erlbaum.

Greiff, S. & Funke, J. (2008). *Indikatoren der Problemlöseleistung: Sinn und Unsinn verschiedener Berechnungsvorschriften. Bericht aus dem MicroDYN Projekt [Measuring Complex Problem Solving: The MicroDYN approach].* Heidelberg: Psychologisches Institut.

Klauer, K. & Phye, G. (2008). Inductive reasoning: a training approach. *Review of Educational Research, 78* (1), 85–123.

Klieme, E. (2004). Assessment of cross-curricular problem-solving competencies. In J.H. Moskowitz and M. Stephens (Eds.). *Comparing Learning Outcomes. International Assessments and Education Policy* (pp. 81–107). London: Routledge Falmer.

Klieme, E., Leutner, D. & Wirth, J. (Eds.). (2005). *Problemlösekompetenz von Schülerinnen und Schülern. Diagnostische Ansätze, theoretische Grundlagen und empirische Befunde der deutschen PISA 2000 Studie* [Problem-solving competency of students. Assessment approaches, theoretical basics, and empirical results of the German PISA 2000 study]. Wiesbaden, Germany: VS Verlag für Sozialwissenschaften.

Lesh, R. & Zawojewski, J. S. (2007). Problem solving and modeling. In F. Lester (Ed.), *The Handbook of Research on Mathematics Teaching and Learning* (2nd ed.) (pp. 763–804). Reston, VA: National Council of Teachers of Mathematics; Charlotte, NC: Information Age Publishing (joint publication).

Leutner, D., Klieme, E., Meyer, K. & Wirth, J. (2004). Problemlösen [Problem solving]. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand, R. Pekrun, J. Rost & U. Schiefele (PISA-Konsortium Deutschland) (Eds.), *PISA 2003: Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs* (pp. 147–175). Münster, Germany: Waxmann.

Leutner, D. & Wirth, J. (2005). What we have learned from PISA so far: a German educational psychology point of view. *KEDI Journal of Educational Policy 2* (2), 39–56.

Mayer, R.E. (1990). Problem solving. In M. W. Eysenck (Ed.), *The Blackwell Dictionary of Cognitive Psychology* (pp. 284–288). Oxford: Basil Blackwell.

Mayer, R. E. (1992). *Thinking, Problem solving, Cognition* (2nd Ed.). New York, NY: Freeman.

Mayer, R. E. (1998). Cognitive, metacognitive, and motivational aspects of problem solving. *Instructional Science 26*, 49–63.

Mayer, R.E. (2002). A taxonomy for computer-based assessment of problem solving. *Computers in Human Behavior 18*, 623–632.

Mayer, R. E. (2003). *Learning and Instruction*. Upper Saddle River, NJ: Merrill Prentice Hall.

Mayer, R. E. & Wittrock, M. C. (1996). Problem-solving transfer. In R. Calfee & R. Berliner (Eds.), *Handbook of Educational Psychology* (pp. 47–62). New York: Macmillan.

Mayer, R. E. & Wittrock, M. C. (2006) Problem Solving. In P. A. Alexander and P. H.Winne (Eds.), *Handbook of Educational Psychology (2nd ed.)* (ch. 13). Mahwah, NJ: Lawrence Erlbaum Associates.

OECD. (2003a). *The PISA 2003 Assessment Framework. Mathematics, Reading, Science and Problem Solving Knowledge and Skills*. Paris: OECD.

OECD. (2003b). The definition and selection of competencies (DeSeCo): Executive summary of the final report. Paris: OECD. http://www.oecd.org/dataoecd/47/61/35070367.pdf

OECD. (2004). *Problem Solving for Tomorrow's World. First Measures of Cross Curricular Competencies from PISA 2003*. Paris: OECD.

OECD. (March 2009). PIAAC problem solving in technology rich environments: Conceptual framework. Paris: OECD.

O'Neil, H. F. (2002). Perspectives on computer-based assessment of problem solving. *Computers in Human Behavior 18*, 605–607.

Osman M. (2010). Controlling uncertainty: A review of human behavior in complex dynamic environments. *Psychological Bulletin 136*, 65–86.

Pólya, G. (1945). *How to Solve It*. Princeton, NJ: Princeton University Press.

Reeff, J.-P., Zabal, A. & Blech. C. (2006). *The Assessment of Problem-Solving Competencies. A Draft Version of a General Framework.* Bonn: Deutsches Institut für Erwachsenenbildung. (Retrieved May 8, 2008, from http://www.die-bonn.de/esprid/dokumente/doc-2006/reeff06_01.pdf).

Robertson, S. I. (2001). *Problem Solving*. East Sussex: Psychology Press.

Rychen D. S. & Salganik, L. H. (Eds.). (2003). *Key Competencies for a Successful Life and a Well-Functioning Society*. Göttingen, Germany: Hogrefe and Huber.

Vosniadou, S. & Ortony, A. (1989). *Similarity and Analogical Reasoning*. New York: Cambridge University Press.

Wirth, J. & Klieme, E. (2004). Computer-based assessment of problem solving competence. *Assessment in Education: Principles, Policy and Practice 10*(3), 329–345.

ANNEX A: OVERVIEW OF PROBLEM SOLVING RESEARCH[1]

Historical and Theoretical Foundations

1. Psychological research on problem solving began in the early 1900s, as an outgrowth of mental philosophy (Humphrey, 1963; Mandler & Mandler, 1964). Throughout the 20th century four theoretical approaches developed: early conceptions, associationism, Gestalt psychology, and information processing (Mayer, in press).

Early Conceptions

2. The start of psychology as a science can be set at 1879 – the year Wilhelm Wundt opened the first world's psychology laboratory in Leipzig, Germany, and sought to train the world's first cohort of experimental psychologists. Wundt's main contribution to the study of problem solving, however, was to call for its banishment on the grounds that complex cognitive processing was too complicated to be studied by experimental methods (Wundt, 1911/1973). In spite of his admonishments, a group of his former students – who came to be known as the Wurzburg group – sought to study thinking by asking students to describe their thought process as they solved word association problems, such as finding a superordinate of "newspaper" (*e.g.*, an answer is "publication"). Although they did not produce a new theoretical approach, the Wurzburg group found empirical evidence that challenged some of the key assumptions of mental philosophy, including challenges to the idea that all thinking involves mental imagery.

Associationism

3. During the 1920s through the 1950s, the first major theoretical approach to take hold in the scientific study of problem solving was *associationism* – the idea that cognitive representations in the mind consist of ideas and links between them and that cognitive processing in the mind involves following a chain of associations from one idea to the next (Mandler & Mandler, 1964; Mayer, 1992). For example, in a classic study, Thorndike (1911) placed a hungry cat in what he called a *puzzle box* – a wooden crate in which pulling a loop of string that hung from overhead would open a trap door to allow the cat to escape to a bowl of food outside the crate. According to Thorndike, over the course of many trials, the cat learned to

[1] Except for the material pertaining to complex problem solving, this review was prepared by Rich Mayer in February 2010. The section on Complex Problem Solving is mainly derived from Funke and Frensch (2007). Professors Mayer (University of California, Santa Barbara) and Funke (University of Heidelberg, Germany) are members of the PISA 2012 Problem Solving Expert Group.

escape through a process that he called the *law of effect*: responses that lead to dissatisfaction become less associated with the situation and responses that lead to satisfaction become more associated with the situation. Solving a problem is simply a matter of trial and error and accidental success, according to Thorndike's associationist view. A major challenge to associationist theory concerns the nature of transfer – that is, explaining how a problem solver invents a creative solution that has never been performed before. Associationist conceptions of cognition can be seen in current research, including neural networks, connectionist models, and parallel distributed processing models (Rogers & McClelland, 2004).

Gestalt Psychology

4. The Gestalt approach to problem solving developed in the 1930s and 1940s as a counterbalance to the associationist approach. According to the Gestalt approach cognitive representations consist of coherent structures (rather than individual associations) and the cognitive process of problem solving involves building a coherent structure (rather than strengthening and weakening of associations). A major focus is on the nature of insight – how problem solvers move from not knowing how to solve a problem to knowing how to solve it (Duncker, 1945; Mayer, 1995). For example, in a classic study, Kohler (1925) observed how a hungry ape in a play yard figured out how to stack crates to allow him to reach a banana hanging overhead. Kohler referred to the underlying mechanism as *insight* – literally seeing into the structure of the situation. A major challenge to Gestalt theory is its lack of precision. Gestalt conceptions can be seen in modern research on mental models and schemas (Gentner & Stevens, 1983).

George Polya

5. George Polya was born in Hungary in 1887 and emigrated to the United States in 1940. A distinguished mathematician, he quickly became known for his research and teachings in problem solving at Stanford University. In his class books, *How to Solve It* (1957), first published in 1945, and *Mathematical Discovery* (1965), Polya broke problem solving into four phases: (1) understanding the problem, (2) devising a plan, (3) carrying out the plan, and (4) looking back. An important contribution of Polya's work is his observation that problem solving is a learnable skill: "you can learn it only by imitation and practice" (Polya, 1965, p. ix). Drawing mainly on examples in mathematical problem solving, Polya suggested numerous problem solving heuristics including thinking of a related problem, breaking a problem into parts, and restating the givens or goal.

Information Processing

6. The information processing approach to problem solving developed in the 1960s and 1970s and was based on the computer metaphor – the idea that humans are processors of information (Mayer, 2009). According to the information processing approach, problem solving involves a series of mental computations – each of which consists of applying a process to a mental representation (such as comparing two elements to determine if they differ). In their classic book, *Human Problem Solving*, Newell and Simon (1972) proposed that problem solving involved a *problem space* – a representation of all the intervening states between the given and goal states – and *search heuristics* – strategies for moving from the given to the goal state. Newell and Simon used computer simulation as a research method to test their conception of human problem solving, in which they built computer programs that solved problems using human-like strategies. An important advantage of the information processing approach is that problem solving can be described with great clarity – as a computer program. An important limitation of the information processing approach is that it is most useful for describing problem solving for well-defined problems rather than ill-defined problems. The information processing conception of cognition lives on as a keystone of today's cognitive science (Mayer, 2009).

Current Lines of Research on Problem Solving

7. The study of problem solving has a somewhat fragmented tradition in which various topics have been studied in relative isolation from one another. Some major lines of problem solving research include decision making, reasoning, intelligence and creativity, teaching of thinking skills, expert problem solving, thinking by analogy, mathematical and scientific problem solving, situated cognition, cognitive neuroscience and complex problem solving.

8. A unifying theme of several of these diverse lines of problem solving research is the idea that problem-solving performance depends on the problem solver's knowledge, including domain-specific knowledge (Mayer & Wittrock, 2006). Another unifying theme concerns the role of the human information processing system – including severe limits on working memory capacity.

Decision Making

9. *Decision making* refers to cognitive processing involved in choosing between two or more alternatives (Barron, 2000; Kahneman & Tversky, 2000; Markman &

Medin, 2002). For example, a decision making task is to choose between (a) an 85% chance of winning $1000 (with a 15% chance of winning nothing) and (b) receiving $800 for sure.

10. Research on decision making has produced a progression of four theories – prescriptive theories, descriptive theories, heuristic theories, and constructive theories. Prescriptive theories – sometimes called economic theories – specify what people should do if they are completely rational. Prescriptive theories – such as expected value theory – predict that people should opt for the first alternative because it has a higher expected value (*i.e.*, $850) than the second alternative (*i.e.*, $800); however, psychological research shows that people prefer the second alternative (Kahneman & Tversky, 1984). Descriptive theories describe what people actually do when confronted with decision tasks. Descriptive theories – such as *prospect theory* – are based on the idea that people prefer to avoid a loss more than they prefer to make an equivalent gain, and that people tend to overestimate small probabilities and underestimate large probabilities. For example most people prefer a sure win of $3000 to an 80% chance of winning $4000, but most people prefer an 80% chance of losing $4000 to a sure loss of $3000 (Kahneman & Tversky, 1979). Heuristic theories build on these findings by focusing on the domain-specific strategies that people use to make decisions (Gigerenzer et al., 1999). For example, most people decide that the letter "r" is more likely to occur in the first position of an English word than in the third position of a word (Tversky & Kahneman, 1973). An explanation is that they are using the *availability heuristic*, in which it is easier to think of words that start with "r" than to think of words that have "r" as the third letter. Constructive theories also build on these findings by focusing on decision-making as a process of constructing a coherent mental model of the situation. For example, people are likely to say they would buy a new $10 theatre ticket if they discover that they had lost $10 but are not likely to buy a new $10 ticket if they discover they had lost the ticket they had already purchased (Kahneman & Tversky, 1984). An explanation for this finding is that people engage in *mental accounting* by which the ticket money was lost in the second scenario but not the first.

11. An important theme in current evidence-based theories of decision making concerns the central role of domain-specific knowledge and cognitive processing that people bring to the decision making task. Another central theme is that people use shortcut strategies to overcome the limits on information processing imposed

by the human information system, thereby relying on what Simon (1982) calls *satisficing* – selecting an acceptable choice that meets one's criteria.

Reasoning

12. *Reasoning* refers to determining whether a conclusion follows from premises (Evans, 2005; Johnson-Laird, 2005). For example, conditional reasoning may be examined using a card-turning problem such as, "If a card has a vowel on one side, it has an even number on the other side. Select those cards that you definitely need to turn over to find out whether or not they violate the rule: A D 4 7" (Wason, 1966). Research shows that people perform better when the card-turning problem is presented in concrete form, such as, "If a person is drinking beer, then the person must be over 19 years of age. Select the card or cards you would definitely need to turn over to determine whether or not the people are violating the rule: DRINKING A BEER, DRINKING A COKE, 16 YEARS OF AGE, 22 YEARS OF AGE" (Griggs & Cox, 1982). Subsequent research has shown that people perform better when the theme of a concrete problem involves detecting cheaters (Cosmides, 1989) or seeking permission (Cheng & Holyoak, 1985). Based on a series of empirical investigations, Johnson-Laird's (2005) mental model theory provides a similar account of deductive reasoning in which people build concrete mental models of the situation described in the premises. These findings point to the domain specificity of reasoning, that is, the idea that people use the concrete context of the problem to guide their reasoning rather than apply the same general logical rules to all problems.

Intelligence and Creativity

13. Research on *intelligence and creativity* explores the role of individual differences in human cognitive ability in problem solving (Guilford, 1967; Sternberg, 1990, 1999; Sternberg & Grigorenko, 2003). Psychometric research examines the correlations among cognitive tests (Carroll, 1993) and evidence for general intelligence (or what has been called *g*) would be reflected in high correlations among all cognitive tests. However, factor analyses (and similar statistical tools) reveal many smaller specialised factors, such as forms of verbal ability, mathematical ability, and spatial ability (Carroll, 1993; Sternberg, 1999). Modern cognitive science research focuses on identifying component processes that support performance on intelligence tests, including cognitive task analyses of intelligence test items; and pinpointing individual differences in the information processing system related to cognitive test performance. This work highlights the role of domain specific knowledge and

processing in cognitive performance, as well as the ways people's thinking is dependent on the architecture of the human information processing system.

Teaching of Thinking Skills

14. Research on *teaching of thinking skills* focuses on training that helps people become better problem solvers (Bloom & Broder, 1950; Covington, Crutchfield, Davies, & Olton, 1974; Nickerson, 1999; Ritchhart & Perkins, 2005). For example, Bloom and Broder (1950) taught college students to solve economics problems by having them watch a model solve problems while thinking aloud. Similarly, Covington, Crutchfield Davies, and Olton (1974) taught elementary school children to solve mystery problems by asking them to read comic books that modelled how to generate and test hypotheses for various mystery problems. A persistent finding is that problem solving training tends to be effective when it focuses on specific skills needed for the task (such as generating and testing hypotheses), when a successful problem solver models the skills, and when the test involves problems that are similar to those used during training. There is not strong evidence that people can be taught general problem solving strategies that improve performance across a diverse set of problem situations (Nickerson, 1999; Ritchhart & Perkins, 2005). Thus, an important theme of research on thinking skills training concerns the role of domain-specific knowledge in supporting problem solving.

Expert Problem Solving

15. Research on *expert problem* solving involves comparing how experts and novices solve problems (Chase & Simon, 1973; de Groot, 1965; Ericsson, Feltovich, & Hoffman, 2006). For example, expert chess players are better able than novices to remember the position of chess pieces on a chessboard from a real game but the difference between experts and novices disappears when the chess pieces are randomly placed on the board (Chase & Simon, 1973). Similarly, when given a chessboard from a real game and asked to make the next move, experts do not consider more moves than novices but experts consider better moves than novices (de Groot, 1965). Research in the development of expertise shows that experts must engage in approximately ten years of practice to gain the knowledge base needed for expert problem solving, but that the knowledge needed for success in one domain does not easily transfer to another (Ericsson, Feltovich, & Hoffman, 2006).

16. These finding are consistent with the idea that experts do not possess better general cognitive skills than novices – such as better memory capacity – but rather

experts have better domain knowledge based on their experience (such as schemas for how several pieces form a single configuration). Similarly, these results also demonstrate the limits on the human information processing system, which necessitates strategies such as mentally arranging the pieces into configurations.

Thinking by Analogy

17. *Thinking by analogy* refers to solving a problem by using one's knowledge about a similar problem (Holyoak, 2005). For example, Gick and Holyoak (1980, 1983) asked students to solve Duncker's (1945) tumour problem, in which they must figure out how to destroy an inoperable tumour by using rays that at sufficient intensity can destroy tissue in their path. Students generally were not successful in producing the desired solution – to have many weak rays all converge on the tumour – even if they were previously given a passage about a general attacking a fortress that was based on the same general solution of having many small forces all converge on the fortress. A major conclusion is that people have trouble engaging in analogical transfer – applying the solution method from one problem that they know to a new problem that can be solved by the same method – particularly when the cover stories differ. This finding also suggests that human cognition tends to be domain specific so that solution strategies used in one context are not automatically transferred to a new context.

Mathematical and Scientific Problem Solving

18. Research on *mathematical and scientific problem solving* focuses on how students solve problems in the specific subject domain of mathematics or science (Mayer, 2008). Research shows that proficiency in solving arithmetic word problems, as an example, depends on specific kinds of knowledge in the problem solver, including facts, concepts, procedures, strategies, and beliefs (Anderson et al., 2001; Kilpatrick, Swafford, & Findell, 2001; Mayer, 2008). For example, students are able to sort arithmetic word problems into categories, suggesting that they possess schemas for specific problem types (Hinsley, Hays, & Simon, 1977). Riley, Greeno, & Heller (1982) found that young students are able to solve word problems presented in a change situation (*e.g.*, John has 3 marbles. He gets 2 more marbles. How many does he have now?) more easily than in a compare situation (*e.g.*, John has 3 marbles. Pete has 2 more marbles than John. How many marbles does Pete have?). Older students performed well on both problem types. This finding

suggests that students may begin with one simple schema and develop additional ones with more experience.

19. In the field of scientific problem solving, a persistent finding is that students enter the science classroom with a set of pre-existing conceptions of how things work, which can be called *preconceptions* or *misconceptions*. For example, many physics students believe that if an object is in motion some force must act on it to keep it in motion – a misconception sometimes referred to as *impetus theory* (McCloskey, 1983). Such misconceptions affect how students make predictions and observe results in a scientific problem solving task, often resulting in *confirmation bias* – attempts to prove one's hypothesis and ignore discrepant data (Chinn & Malhotra, 2002; Dunbar, 1993). Overall, research on mathematical and scientific problem solving points to the role of the learner's specific knowledge in determining the process of problem solving.

Situated Cognition

20. *Situated cognition* refers to solving problems within specific physical, social and cultural contexts (Nunes, Schlieman, & Carraher, 1993; Robbins & Aydede, 2009). For example, in a classic study, Nunes, Schlieman, and Carraher (1993) found that Brazilian students used completely different computational strategies for solving arithmetic problems that they received on paper within a school versus problems they encountered in their job as street vendors. In school, when given a problem such as 35 x 10 = ___ they tried to apply the school-taught procedures for multiplication, whereas in the street they invented a strategy based on repeated addition (105 + 105 + 105 + 35 = 350). In a study of computational strategies used by grocery store shoppers to determine which of two items was better to buy, Lave (1988) found that people never used the school-taught strategy of computing unit cost and instead used different kinds of strategies for different situations. For example, to choose between a 10-ounce can of peanuts for 90 cents and a 4-ounce can for 45 cents, they used a ratio strategy in which the larger can is a better buy because it costs twice as much but has more than twice as many ounces. Research on situated cognition – or what can be called *everyday thinking* – shows that the context in which someone encounters a problem influences how they go about solving the problem.

Cognitive Neuroscience of Problem Solving

21. Research on the *cognitive neuroscience of problem solving* focuses on the brain activity that occurs during problem solving (Goel, 2005). For example, Goel (2005)

found different parts of the brain are recruited for solving reasoning problems that are presented in abstract form (*e.g.*, “All P are B. All C are P. Therefore, all C are B.”) versus in concrete form (*e.g.*, “All dogs are pets. All poodles are dogs. Therefore, all poodles are pets.”). Such findings again suggest that reasoning is somewhat domain specific rather than being based on applying a universal set of reasoning rules.

Complex Problem Solving

22. About 30 years ago, Dörner (1975) adopted the term *complex problem solving* to describe a particular type of problem to be studied, which differed from simple problem solving in terms of complexity and other attributes, as well as a new method of investigation, the use of computer-simulated microworlds. At this time, researchers were becoming increasingly convinced that empirical findings and theoretical concepts derived from simple novel tasks could not easily be generalised to more complex, real-life problems, and that domain-specific knowledge and strategies play an important part in problem solving.

23. In Europe, two main approaches to complex problem solving research emerged having in common an emphasis on relatively common, semantically rich, computerised laboratory tasks constructed to be similar to real-life problems (Frensch & Funke, 1995; Funke & Frensch, 2007). One approach, initiated by Broadbent (1977; see Berry & Broadbent, 1995), focuses on the distinction between *explicit problem solving* (problem solving controlled by the solver’s intentions) and *implicit problem solving* (problem solving that is automatic or nonconscious), and typically employs mathematically well-defined computer models. The other approach, initiated by Dörner (*e.g.,* Dörner, 1980; Dörner et al, 1983), is “interested in the interplay of the cognitive, motivational, and social components of problem solving” (Funke & Frensch, 2007), and utilises very complex computer simulations involving up to 2000 interconnected variables.

24. Frensch and Funke (1995b, p. 18) propose the following distinctive characteristics of complex problem solving in keeping with the research traditions established by Broadbent and Dörner[2]: “*The given state, goal state, and barriers between given state and goal state are complex, change dynamically during problem solving, and*

[2] This definition is rooted in the European tradition of Broadbent and Dörner. A much broader view of what constitutes a complex problem, in contrast to a simple problem, is taken by many authors (*e.g.,* Sternberg & Frensch, 1991). Some use it to denote authentic tasks, irrespective of their properties; others use it in a similar manner to Frensch and Funke, but allow complex problems to be dynamic *or* intransparent.

are intransparent. The exact properties of the given state, goal state, and barriers are unknown to the solver at the outset". These features can be illustrated with an example (due to Vollmeyer, Burns & Holyoak, 1996). A computer-simulated fish tank is presented to subjects. In the simulation there are four species of sea animal (crabs, prawns, lobsters and sea bass); these species are affected by four quantities (temperature, salt, oxygen and current). The four quantities can be manipulated by the subjects (users of the system), but the relationships between the quantities and the sea animals are not given to subjects. Changes made by subjects to temperature, salt, oxygen and current will cause changes in the sea animal populations. In addition, in the absence of any such subject intervention the lobster population will decline *of its own accord*. Subjects are asked to try to work out the relationships between the quantities and the sea animals by experimenting with different input values, and are then asked to control the system in order to reach specific target values of each sea animal. It should be noted that the equations governing the underlying system are quite arbitrary, implying that knowledge of real-world marine biology will be at best useless and at worst counter-productive.

25. Research (Funke & Frensch, 2007) does not support a strong link between global intelligence and complex problem solving competency when goal specificity and transparency are low and when the semantic content is rich, but some intelligence components (in particular, processing capacity-reasoning ability and learning potential) appear to be correlated with complex problem solving competency even when the task is intransparent. However, as is the case with simple problems, complex problem solving competency is highly dependent on domain-specific knowledge and strategies. Finally, whether complex problem solving competency is an independent construct to "simple problem solving" is an open question though the results of the findings of the German national extensions of PISA 2000 (Wirth & Klieme, 2004) and PISA 2003 (Leutner *et al*, 2004; Leutner & Wirth, 2005) give some support to the existence of an independent complex problem solving construct.

Conclusions

26. A major implication of problem-solving research for the PISA 2012 Problem Solving assessment is that it should take into account the role of the student's domain knowledge in problem solving rather than trying to assess problem solving in general. In addition, the problem-solving literature suggests that the PISA 2012 Problem Solving assessment should be based on problem solving in authentic concrete contexts rather than with abstract tasks. The problem-solving literature

also suggests that the problem-solving tasks used in the PISA 2012 Problem Solving assessment should be sufficiently rich to allow for students to use their information processing systems to manage a collection of different kinds of their knowledge rather than be based on short puzzle-like items.

27. Increasingly, the problems facing citizens of the 21st century are complex problems and methods for solving simple problems cannot be easily generalised to solve such problems. Furthermore, there is no empirical evidence for a relation between complex problem solving competency and global intelligence, but there is some evidence (though not conclusive) that complex problem solving competency is a separate construct and not just the application of “normal” problem-solving processes to complex situations. The clear implication is that complex problem solving should be a central feature of the PISA 2012 Problem Solving assessment.

References

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.* New York: Longman.

Baron, J. (2000). *Thinking and deciding* (3rd ed). New York: Cambridge University Press.

Bloom, B. S., & Broder, B. J. (1950). *Problem-solving processes of college students: An exploratory investigation.* Chicago: University of Chicago Press.

Berry, D. C., & Broadbent, D. E. (1995). Implicit learning in the control of complex systems. In P.A. Frensch & J. Funke (Eds.), *Complex Problem Solving: The European Perspective* (pp. 3–25). Hillsdale, NJ: Lawrence Erlbaum Associates.

Broadbent, D. E. (1977). Levels, hierarchies, and the locus of control. *Quarterly Journal of Experimental Psychology 29*, 181–200.

Carroll, J. B. (1993). *Human cognitive ability.* New York: Cambridge University Press.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*, 55–81.

Cheng, P. W., & Holyoak, K. (1985). Pragmatic reasoning schemas. *Cognitive Psychology, 17,* 391–416.

Chinn, C. A., & Malhotra, B. A. (2002). Children's responses to anomalous scientific data: How is conceptual change impeded? *Journal of Educational Psychology, 94,* 327–343.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? *Cognition, 31,* 187–276.

Covington, M. V., Crutchfield, R. S., Davies, L. B., & Olton, R. M. (1974). *The productive thinking program.* Columbus, OH: Merrill.

de Groot, A. D. (1965). *Thought and choice in chess.* The Hague, Netherlands: Mouton.

Dörner, D., Kreuzig, H.W., Reither, F., & Stäudel, T. (1983). *Lohhausen. Vom Unigang mit Unbestimmtheit und Komplexität* [*Lohhausen. On dealing with uncertainty and complexity*].Bern: Huber.

Dorner, D. (1980). On the difficulty people have in dealing with complexity. *Simulation & Games 11*, 87–106.

Dienes, Z., & Berry, D. (1997). Implicit learning: Below the subjective threshold. *Psychonomic Bulletin & Review 4*, 3–23.

Dunbar, K. (1993). Concept discovery in a scientific domain. *Cognitive Science, 17,* 397–434.

Duncker, K. (1945). On problem solving. *Psychological Monographs 58*:3 (Whole No. 270).

Ericsson, K. A., Feltovich, P. J., & Hoffman, R. R. (Eds.) (2006). *The Cambridge handbook of expertise and expert performance.* New York: Cambridge University Press.

Evans, J. S. B. T. (2005). Deductive reasoning. In K. J. Holyoak, & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 169–184). New York: Cambridge University Press.

Frensch, P. A. & Funke, J. (Eds.). (1995). *Complex Problem Solving: The European Perspective.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Frensch, P.A. & Rünger, D. (2003) Implicit learning. *Current Directions in Psychological Science 12*, 13–18.

Funke, J. (2001). Dynamic systems as tools for analysing human judgement. *Thinking and Reasoning 7*(1)*,* 69–89.

Funke, J. and Frensch, P. A. (2007). Complex problem solving: The European perspective – 10 years after. In D. H. Jonassen (Ed.), *Learning to Solve Complex Scientific Problems* (pp. 25–47). New York: Lawrence Erlbaum.

Gentner, D., & Stevens, A. L. (Eds.). (1983). *Mental models.* Hillsdale, NJ: Erlbaum.

Gick, M. I., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12,* 306–355.

Gick, M. I., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*, 1–38.
Gigerenzer, G., Todd, P. M., & the ABC Research Group (Eds.). (1999). *Simple heuristics that make us smart.* Oxford, UK: Oxford University Press.
Goel, V. (2005). Cognitive neuroscience of deductive reasoning. In K. J. Holyoak, & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 475–492). New York: Cambridge University Press.
Griggs, R. A., & Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology, 73*, 407–420.
Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.
Hinsley, D., Hayes, J. R., & Simon, H. A. (1977). From words to equations. In P. Carpenter & M. Just (Eds.), *Cognitive processes in comprehension*. Hillsdale, NJ: Erlbaum.
Holyoak, K. J. (2005). Analogy. In K. J. Holyoak, & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 117–142). New York: Cambridge University Press.
Humphrey, G. (1963). *Thinking: An introduction to experimental psychology.* New York: Wiley.
Johnson-Laird, P. N. (2005). Mental models and thought. In K. J. Holyoak, & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 185–208). New York: Cambridge University Press.
Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263–291.
Kahneman, D., & Tversky, A. (1984). Choice, values, and frames. *American Psychologist, 39*, 341–350.
Kahneman, D., & Tversky, A. (Eds.). (2000). *Choices, values, and frames*. New York: Cambridge University Press.
Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.
Klieme, E. (2004). Assessment of cross-curricular problem-solving competencies. In J.H. Moskowitz and M. Stephens (Eds.). *Comparing Learning Outcomes. International Assessments and Education Policy* (pp. 81–107). London: Routledge Falmer.
Kohler, W. (1925). *The mentality of apes.* New York: Liveright.
Lave, J. (1988). *Cognition in practice*. Cambridge, England: Cambridge University Press.
Leutner, D., Klieme, E., Meyer, K. & Wirth, J. (2004). Problemlösen [Problem solving]. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand, R. Pekrun, J. Rost & U. Schiefele (PISA-Konsortium Deutschland) (Eds.), *PISA 2003: Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs* (pp. 147–175). Münster, Germany: Waxmann.
Mandler, J. M., & Mandler, G. (1964). *Thinking from associationism to Gestalt*. New York: Wiley.
Markman, A. B., & Medin, D. L. (2002). Decision making. In D. Medin (Ed.), *Stevens' handbook of experimental psychology, Volume 2: Memory and cognitive processes* (2nd ed; pp. 413–466). New York: Wiley.
Mayer, R. E. (1992). *Thinking, problem solving, cognition* (2nd ed). New York: Freeman.
Mayer, R. E. (1995). The search for insight: Grappling with Gestalt psychology's unanswered questions. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 3–32). Cambridge, MA: MIT Press.
Mayer, R. E. (2008). *Learning and instruction.* Upper Saddle River, NJ: Merrill Prentice Hall.
Mayer, R. E. (2009). Information processing. In T. L. Good (Ed.), *21st century education: A reference handbook* (pp. 168–174). Thousand Oaks, CA: SAGE.
Mayer, R. E. (in press). Problem solving. In D. Reisberg (Ed.), *Oxford handbook of cognitive psychology*. New York: Oxford University Press.

Mayer, R. E., & Wittrock, M. C. (2006). Problem solving. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed; pp. 287–304). Mahwah, NJ: Erlbaum.

McCloskey, M. (1983). Intuitive physics. *Scientific American, 248(*4), 122–130.

Newell, A., & Simon, H. A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice-Hall.

Nickerson, R. S. (1999). Enhancing creativity. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 392–430). New York: Cambridge University Press.

Nunes, T., Schliemann, A. D., & Carraher, D. W, (1993*). Street mathematics and school mathematics.* Cambridge, UK: Cambridge University Press.

Polya, G. (1957). *How to solve it*. Garden City, NY: Doubleday. [Originally published in 1945 by Princeton University Press.]

Polya, G. (1965). *Mathematical discovery (vol. 2)*. New York: Wiley.

Osman M. (2010). Controlling uncertainty: A review of human behavior in complex dynamic environments. *Psychological Bulletin 136,* 65–86.

Riley, M., Greeno, J. G., & Heller, J. (1982). The development of children's problem solving ability in arithmetic. In H. Ginsberg (Ed.), *The development of mathematical thinking* (pp. 153–199). New York: Academic Press.

Ritchhart, R., & Perkins, D. N. (2005). Learning to think: The challenge of teaching thinking. In K. J. Holyoak, & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 775–802). New York: Cambridge University Press.

Robbins, P., & Aydede, M. (Eds.). (2009). *The Cambridge handbook of situated cognition.* New York: Cambridge University Press.

Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach.* Cambridge, MA: MIT Press.

Simon, H. A. (1982). *Models of bounded rationality.* Cambridge, MA: MIT Press.

Sternberg, R. J. (1990). *Metaphors of mind: Conceptions of the nature of intelligence.* New York: Cambridge University Press.

Sternberg, R. J. (1999). *Handbook of creativity.* New York: Cambridge University Press.

Sternberg, R. J., & Frensch, P. A. (Eds.). (1991). *Complex problem solving: Principles and mechanisms*. Hillsdale, NJ: Lawrence Erlbaum Associates

Sternberg, R. J., & Grigorenko, E. L. (Eds.). (2003). *The psychology of abilities, competencies, and expertise.* New York: Cambridge University Press.

Thorndike, E. L. (1911). *Animal intelligence.* New York: Hafner.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*, 207–232.

Vollmeyer, R., Burns, B. D., & Holyoak, K. J. (1996). The impact of goal specificity and systematicity of strategies on the acquisition of problem structure. *Cognitive Science 20*, 75–100.

Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology.* Harmondsworth, England: Penguin.

Wirth, J. and Klieme, E. (2004). Computer-based assessment of problem solving competence. *Assessment in Education: Principles, Policy and Practice 10*(3), 329–345.

Wundt, W. (1973). *An introduction to experimental psychology.* New York: Arno Press. [Originally published in German in 1911.]

ANNEX B: PROBLEM SOLVING EXPERT GROUP

The PEG membership for PISA 2012 is as follows:

PEG Member	Affiliation
Joachim Funke (Chair)	Heidelberg University, Germany
Benő Csapó	University of Szeged, Hungary (*Ex officio* PGB representative)
John Dossey	Illinois State University, USA
Art Graesser	University of Memphis, USA
Detlev Leutner	Duisburg-Essen University, Germany
Richard Mayer	University of California, USA
Tan Ming Ming	Ministry of Education, Singapore
Romain Martin	University of Luxembourg, Luxembourg