7

PISA 2015 collaborative problem-solving framework

This chapter describes the rationale behind measuring 15-year-olds' collaborative problem-solving skills for the first time in the Programme for International Student Assessment (PISA). It explains the content and processes that are reflected in the collaborative problem-solving items used in the computer-based assessment, and describes how student proficiency in this domain is measured and reported.

Collaborative problem solving (CPS) is a critical and necessary skill used in education and in the workforce. While problem solving as defined in PISA 2012 (OECD, 2010) relates to individuals working alone on resolving problems where a method of solution is not immediately obvious, in CPS, individuals pool their understanding and effort and work together to solve these problems. Collaboration has distinct advantages over individual problem solving because it allows for:

- an effective division of labour
- the incorporation of information from multiple perspectives, experiences and sources of knowledge
- enhanced creativity and quality of solutions stimulated by the ideas of other group members.

Collaboration has been defined as a "co-ordinated, synchronous activity that is the result of a continued attempt to construct and maintain a shared conception of a problem" (Roschelle and Teasley, 1995, p. 70). Social interaction is a vital but insufficient condition for collaboration because some social interactions do not involve shared goals, accommodation of different perspectives or organised attempts to achieve the goals.

There is a growing emphasis in state and national education systems on project-based and inquiry-oriented learning (National Research Council, 2011a). This includes shaping curriculum and instruction around critical thinking, problem solving, self-management and collaboration skills (Darling-Hammond, 2011; Halpern, 2003). Project-based work often includes tasks that require multiple students working together to achieve a team goal, such as a final report, integrated analyses or a joint presentation. Collaborative problem solving is not typically taught as an independent skill distinct from particular subjects. Therefore, in school-based contexts, collaborative learning exercises are often integrated into specific courses of study, such as the sciences, mathematics and history.

Recent curriculum and instruction reforms have focused to a greater extent on the teaching and assessment of 21st century skills (Griffin, Care and McGaw, 2011; National Research Council, 2011a,b). These skills have included critical thinking, problem solving, self-management, information and communication technology (ICT) skills, communication and collaboration (Binkley at al., 2011; OECD, 2011). Collaboration and communication skills are central to this 21st-century skill set and are described in a number of 21st-century skills curricula and assessment reports.

For example, the focal point of Singapore's third IT Masterplan (2009-14) is to facilitate a greater level of technological integration in curriculum, assessment and pedagogy in order to equip students with critical competencies, such as self-directed learning and collaboration skills (Ministry of Education, Singapore, 2008). Similarly, the Israeli national programme, Adapting the Educational System to the 21st Century (Ministry of Education, 2011), is a multiple-year plan with the goal of introducing innovative pedagogy in schools, including communication, collaboration, and other 21st-century skills. However, many of these curricula provide only a general framework and a description of goals and curriculum standards without defining the specific collaboration skills that are to be taught (Darling-Hammond, 2011).

Students need to prepare for careers that require the ability to work effectively in groups and to apply their problem-solving skills in these social situations (Brannick and Prince, 1997; Griffin, Care and McGaw, 2011; National Research Council, 2011a; Rosen and Rimor, 2012). There has been a marked shift from manufacturing to information and knowledge services. Much of the problem-solving work carried out in the world today is performed by teams in an increasingly global and computerised economy. However, even in manufacturing, work is seldom conducted by individuals working alone. Moreover, with the greater availability of networked computers, individuals are increasingly expected to work with diverse teams spread across different locations using collaborative technology (Kanter, 1994; Salas, Cooke and Rosen, 2008).

The University of Phoenix Research Institute identified virtual collaboration, i.e. the "ability to work productively, drive engagement, and demonstrate presence as a member of a virtual team" (Davis, Fidler and Gorbis, 2011, p. 12), as one of ten key skills for the future workforce. A recent Forrester report, based on a survey of information and knowledge-management decision makers from 921 North American and European enterprises, revealed that 94% had implemented or were going to implement some form of collaboration technologies, including e-mail, web conferencing, team workspaces, instant messaging or videoconferencing (Enterprise and SMB Software Survey, North America and Europe, Q42009 Forrester report). CPS skills are also needed in civic contexts, such as social networking, volunteering, participation in community life, and transactions with administration and public services. Thus, students emerging from schools into the workforce and public life will be expected to have collaborative problem-solving skills and the ability to collaborate using appropriate technology.

Collaboration among team members is crucial to the success of groups, families, corporations, public institutions, organisations and government agencies. One uncooperative member of a team can have serious negative consequences on team success. Skilled collaboration and social communication facilitate performance in the workplace (Klein, DeRouin and Salas, 2006; Salas, Cooke and Rosen, 2008), in engineering and software development (Sonnentag and Lange, 2002), and in interdisciplinary research among scientists (Nash et al., 2003). This is clearly apparent from the trend in research publications. Wuchty, Jones and Uzzi (2007) examined 19.9 million papers over five decades and demonstrated that there has been an increase in publications by teams of authors. Moreover, papers drafted by teams of authors end up higher in citation indices than papers drafted by individual authors.

The competencies assessed in the PISA 2015 collaborative problem-solving assessment therefore need to reflect the skills found in project-based learning in schools and in collaboration in workplace and civic settings, as described above. In such settings, students are expected to be proficient in skills such as communicating, managing conflict, organising a team, building consensus and managing progress.

One major factor that contributes to the success of CPS is effective communication among team members (Dillenbourg and Traum, 2006; Fiore et al., 2010; Fiore and Schooler, 2004). Therefore, an important part of the assessment must be proficiency in communication: communicating the right information and reporting what actions have been taken to the right person at the right time. This allows students to build a shared understanding of the task. The competency includes considering the perspectives of other team members, tracking the knowledge of team members, and building and monitoring a shared understanding of the progress made on the task.

Students must also be able to establish and maintain effective team organisation. This includes understanding and assigning roles, and maintaining and adapting the organisation to be effective at achieving its goals. This includes handling disagreements, conflicts, obstacles to goals and potential negative emotions (Barth and Funke, 2010; Dillenbourg, 1999; Rosen and Rimor, 2009).

In addition, students need to understand the type of collaboration and associated rules of engagement. The ground rules are different in contexts of helping, collaborative work, consensus building, win-win negotiations, debates and hidden-profile jigsaw configurations (i.e. group members have different information that needs to be integrated to arrive at a solution).

Apart from defining the domain, the CPS framework has to propose a way to operationalise the construct through a computer-based assessment (CBA). The framework builds, in part, on the individual problem-solving framework for PISA 2012, but extends it substantially in order to cover the additional concepts that need to be incorporated in order to develop and focus on the collaborative aspects of problem solving. The main elements of these aspects are group thinking and the communication skills required for effective interaction between group and individual thinking.

The CPS framework incorporates definitions and theoretical constructs that are based on research and best practices from several areas where CPS-related skills have been assessed. These areas include computer-supported co-operative work, team-discourse analysis, knowledge sharing, individual problem solving, organisational psychology, and assessment in work contexts (e.g. military teams, corporate leadership). The framework further incorporates information from existing assessments that can inform the PISA 2015 CPS assessment, including Assessment and Teaching of 21st-Century Skills (ATC21s), problem solving in the Programme for the International Assessment of Adult Competencies (PIAAC), Partnership for 21st-Century Skills, and the PISA 2012 individual problem-solving assessment (see Annex 7.B for a review).

The operationalisation of the framework described in section four requires an understanding of the major theoretical and logistical underpinnings of an assessment. The framework cannot be developed independently of considerations of the assessment design and measurement requirements. It must take into account the types of technologies, tasks and assessment contexts in which it will be applied (Funke, 1998; Funke and Frensch, 2007). For assessment design, the framework must consider the kinds of constructs that can be reliably measured, and must provide valid inferences about the collaborative skills being measured and about their impact on success in today's world. The CPS framework must also provide a basis for the development of computer-based assessments that will be used worldwide within the logistical constraints and time limits of an international assessment.

This document is organised into four sections. Following this introductory section, the section "Defining the domain" provides a definition of collaborative problem solving. The section "Organisation of the domain" describes how the domain of CPS is organised. It explains the skills and competencies needed for successful CPS and the factors that influence these skills. The section "Assessing collaborative problem-solving competency" operationalises the construct

of CPS by identifying and justifying approaches to measuring CPS competencies and the contexts in which the skills can be assessed. It also describes the levels of proficiency for CPS and how they are reported. Annex 7.A provides a summary of studies with conversational agents in tasks that involve tutoring, collaborative learning, co-construction of knowledge, and collaborative problem solving. Annex 7.B provides a literature review of the key concepts in CPS related to the definition, constructs and design decisions of PISA 2015 CPS framework. Annex 7.C provides two CPS units that were developed as preliminary samples to illustrate the assessment framework and show how it might be operationalised.

DEFINING THE DOMAIN

Collaborative problem solving

The PISA 2003 Assessment Framework: Mathematics, Reading, Science and Problem Solving Knowledge and Skills (OECD, 2003) defines problem-solving competencies as:

> *… 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 content areas or curricular areas that might be applicable are not within a single subject area of mathematics, science or reading.*

The draft framework for the individual problem-solving domain in PISA 2012 (OECD, 2010) largely reiterates the 2003 definition but adds an affective element:

> *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.*

In defining the domain of collaborative problem solving for PISA 2015, the aspect of collaboration is obviously the most salient addition to previous versions of the domain of problem solving in PISA. In the definition for the 2015 domain, the emphasis is therefore on this collaborative aspect. The definition identifies the main elements of the domain and the relationships among these elements.

For the purposes of the assessment, the PISA 2015 definition of CPS competency is articulated in Box 7.1.

> Box 7.1. **Definition of collaborative problem solving for PISA 2015**
>
> Collaborative problem-solving competency is the capacity of an individual to effectively engage in a process whereby two or more agents attempt to solve a problem by sharing the understanding and effort required to come to a solution and pooling their knowledge, skills and efforts to reach that solution.

PISA 2015 CPS competency is a conjoint dimension of collaboration skills and the skills needed to solve the problem (i.e. referential problem-solving skills), while collaboration serves as a leading strand.

The following remarks are offered to clarify the meaning and use of the constituent elements of the definition given above.

The capacity of an individual…

Collaboration skills can be assessed at the individual, group, or organisational level (Campbell, 1968; Dillenbourg, 1999; Fiore et al., 2010; Stahl, 2006). An advantage of collaboration is that the output of the group in solving the problem can be greater than the sum of the outputs from individual members (Aronson and Patnoe, 1997; Dillenbourg, 1999; Schwartz, 1995) and the individual level of participants does not adequately characterise how the group as a whole produces different outcomes than individuals. Yet, for the purpose of the PISA assessment, the focus is on individual capacities *within* collaborative situations. The effectiveness of collaborative problem solving depends on the ability of group members to collaborate and to prioritise the success of the group over individual successes. At the same time, this ability is a trait in each of the individual members of the group.

…to effectively engage in a process…

Collaborative problem solving involves an individual's cognitive processing that engages both cognitive and social skills. There are individual problem-solving processes as well as communication processes that interact with the cognitive systems of the other participants in the collaboration. For example, the group must not only have the correct solution

but must also agree that it is the correct solution. As discussed later in this document, the focus of the assessment is the cognitive and social skills related to CPS to establish and maintain shared understanding, to take appropriate actions to solve the problems, and to establish and maintain group organisation.

The cognitive processes involved in CPS are internal to the individual but they are also manifested in the interactions with the problem and with others in the group. That is, cognitive processes can be inferred from the actions performed by the individual, communications made to others, intermediate and final products of the problem-solving tasks, and open-ended reflections on problem-solving representations and activities. These measures can be instantiated by examining exploration and solving strategies, the type and quality of communication generated, probes of the knowledge and representation of the problem, and indicators of an individual's representation of others in the group. In other words, measuring collaborative problem-solving skills is not only a challenge comparable to measuring individual skills, but also a great opportunity to make observable the cognitive processes engaged by team members.

...whereby two or more agents ...

Collaboration requires interactions between two or more agents. The word "agent" refers to either a human or a computer-simulated participant. In both cases, an agent has the capability of generating goals, performing actions, communicating messages, reacting to messages from other participants, sensing its environment, adapting to changing environments, and learning (Franklin and Graesser, 1996). The success of CPS skills can be observed at either an individual level or a group level. Even when observations are directed at an individual level, they refer to the individual's actions and interaction enacted in order to share a representation or common goal with at least one other agent for there to be collaboration. The definition therefore sets the requirement of a minimum of two agents.

...attempt to solve a problem...

The measurement is focused primarily on the collaborative actions the students engage in while trying to solve the problem at hand, rather than solely on the correct solution of the problem. The core construct weighs collaboration processes higher than the solutions to problems.

...by sharing the understanding and effort required to come to a solution...

Collaboration can only occur if the group members strive for building and maintaining a shared understanding of the task and its solutions. Shared understanding is achieved by constructing a common ground (Clark, 1996; Clark and Brennan, 1991; Fiore and Schooler, 2004) through communication and interaction, such as building a shared representation of the meaning of the problem, understanding each individual's role, understanding the abilities and perspectives of group members, mutual tracking of the transfer of information and feedback among group members, and mutual monitoring of progress towards the solution.

...and pooling their knowledge, skills and effort to reach that solution.

Collaboration further requires that each individual establish how their own knowledge and skills can contribute to solving the problem as well as identify and appreciate the knowledge and skills that the other participant(s) can contribute. In addition to establishing the state of the pooled knowledge and skills within the group, there are potential differences in points of view, dissension/conflict among group members, errors committed by group members in need of repair, and other challenges in the problem that require cognitive effort to handle. This additional effort of justifying, defending, arguing and reformulating is a factor that may explain why groups sometimes achieve more or are more efficient than individuals: they have to be explicit about their opinion, interpretations and suggestions requiring them to process available information more deeply, to compare more solutions, and to find out the weaknesses of each solution. If there is no effort from an individual, then that individual is not collaborating. The individual is not expending productive effort if the individual does not respond to requests or events and does not take actions that are relevant to any progress towards goals.

ORGANISATION OF THE DOMAIN

Processes and factors affecting collaborative problem solving

Collaborative problem solving is an inherently complex mechanism that incorporates the components of cognition found in individual problem solving in addition to the components of collaboration. The cognitive components of individual problem solving include understanding and representing the problem content, applying problem-solving strategies, and applying self-regulation and metacognitive processes to monitor progress towards the goal (Funke, 2010; Glaser, Linn and

Bohrnstedt, 1997; Hacker, Dunlosky and Graesser, 2009; Mayer, 1998; O'Neil, 1999). However, engaging other group members in a collaborative task requires additional cognitive and social skills to allow shared understanding, knowledge and information flow, to create and understand an appropriate team organisation, and to perform co-ordinated actions to solve the problem (Dillenbourg, 1999; Fiore et al., 2010).

For the purpose of the PISA 2015 CPS assessment, collaborative problem-solving competency is defined in Box 7.1 as the capacity of an individual to effectively engage in a process whereby two or more agents attempt to solve a problem by sharing the understanding and effort required to come to a solution, and pooling their knowledge, skills and effort to reach that solution. The definition incorporates three core collaborative problem-solving competencies:

1. establishing and maintaining shared understanding
2. taking appropriate action to solve the problem
3. establishing and maintaining team organisation.

These three competencies arise from a combination of collaboration and individual problem-solving processes. The individual problem-solving processes are already defined by the PISA 2012 framework: *exploring and understanding; representing and formulating; planning and executing;* and *monitoring and reflecting*. The CPS competencies are further influenced by factors such as the task, the team composition, the medium in which the task is applied, as well as the overall background context of the problem-solving task. Below we elaborate on these components.

Problem-solving skills

Much of the basis and terminology of collaborative problem solving for PISA 2015 is consistent with that of the PISA 2012 individual problem-solving framework, which addressed problem solving by an individual working alone. It defines a *problem* as existing when a person has a goal but does not have an immediate solution as to how to achieve it. That is, "problem solving is the cognitive processing directed at transforming a given situation into a goal situation when no obvious method of solution is available" (Mayer 1990, p. 284). *Problem-solving competency* is defined as "an individual's capacity to engage in cognitive processing to understand and resolve problem situations where a method of solution is not immediately obvious. It includes the willingness to engage with such situations in order to achieve one's potential as a constructive and reflective citizen" (OECD, 2010, p. 12).

The 2012 individual problem-solving framework identifies three conceptual dimensions that provide the basis for the assessment of problem solving and are also relevant to CPS. These are the *problem context*, the *nature of the problem situation*, and the *problem-solving process* (OECD, 2010, p. 16).

The *problem context* affects how difficult a problem will be to solve for individuals who have varying familiarity with the context. The 2012 individual problem-solving framework posits two aspects of the problem-solving context, namely the *setting* (whether or not it is based on technology) and the *focus* (whether it is personal or social). When the setting is based on technology, individual problem solvers make use of a technological device as a context for their problem solving, such as a computer, cell phone or remote control. The typical problem-solving goal in this context is understanding how to control or troubleshoot the device. Other problem-solving contexts do not make use of such devices. The non-technology contexts include route planning, task scheduling, and decision making (OECD, 2010, p. 17). The focus of the problem solving is classified as personal when it relates mainly to the individual being assessed, the person's family, or the person's peers. A social focus, on the other hand, is broader in the sense that it refers to a context in the wider community or society at large.

The *nature of the problem situation* describes whether the information about the problem situation is complete or not when initially presented to the problem solver. Those problem situations that are complete in their information are referred to as *static* problem situations. When it is necessary for the problem solver to explore the problem situation in order to obtain additional information that was not provided at the onset, the problem situation is referred to as *interactive*. Problem situations may also vary with respect to the degree to which the starting state of the problem, the goal state, and the actions that can be performed to achieve the goal state are specified. Problem situations for which there are clearly specified goals, given states, and legal actions can be labelled *well-defined* problems; in contrast, problems that involve multiple goals in conflict with underspecified given states and actions are called *ill-defined* problems. The PISA 2012 problem-solving assessment and the problem solving in technology-rich environments assessment that is part of the Survey of Adult Skills, a product of the OECD Programme for the International Assessment of Adult Competencies (PIAAC), presented both well-defined and ill-defined problems (OECD, 2010, 2009).

The PISA 2012 individual problem-solving framework identified the following four cognitive processes in individual problem solving: *exploring and understanding; representing and formulating; planning and executing;* and *monitoring and reflecting* (OECD, 2010, p. 20-21). Similar processes were also identified in the PIAAC problem solving in technology-rich environments framework, with the latter being more focused on processes related to the acquisition, use and production of information in computerised environments (OECD, 2009). The CPS framework builds on the previous assessments of individual problem solving with these cognitive processes.

The first process involves understanding the problem situation by interpreting initial information about the problem and any information that is uncovered during exploration and interactions with the problem. In the second process, this information is selected, organised, and integrated with prior knowledge. This is accomplished by representing the information using graphs, tables, symbols and words, and then formulating hypotheses by identifying the relevant factors of the problem and critically evaluating information. The third process includes planning, which consists of clarifying the goal of the problem, setting any subgoals, and developing a plan to reach the goal state. Executing the plan that was created is also a part of this process. The final process consists of monitoring steps in the plan to reach the goal state and reflecting on possible solutions and critical assumptions.

These four problem-solving processes provide a basis for the development of the cognitive strand of the conjoint dimension of the CPS framework. In collaborative problem solving, the group must perform these problem-solving processes concurrently with a set of collaborative processes.

Collaborative problem-solving skills and competencies

Three major collaborative problem-solving competencies are identified and defined for measurement in the assessment. These three major CPS *competencies* are crossed with the four major individual problem-solving *processes* to form a matrix of specific *skills*. The specific skills have associated actions, processes and strategies that define what it means for the student to be competent. Table 7.1 outlines the skills of collaborative problem solving as a matrix of these collaborative and individual processes. The matrix incorporates the individual problem-solving processes from the PISA 2012 individual problem-solving framework and illustrates how each interacts with the three collaboration processes.

The CPS skills identified in this framework are based on a review of other CPS frameworks, such as the National Center for Research on Evaluation, Standards and Student Testing (CRESST) teamwork processing model (O'Neil et al., 2010, 2003), the teamwork model of Salas and colleagues (Fiore et al., 2010, 2008; Salas et al., 2008, 1992) and ATC21s (Griffin, Care and McGaw, 2011). Annex 7.B provides a review of related frameworks and CPS research.

Table 7.1 Matrix of collaborative problem-solving skills for PISA 2015

	(1) Establishing and maintaining shared understanding	**(2) Taking appropriate action to solve the problem**	**(3) Establishing and maintaining team organisation**
(A) **Exploring and understanding**	**(A1)** Discovering perspectives and abilities of team members	**(A2)** Discovering the type of collaborative interaction to solve the problem, along with goals	**(A3)** Understanding roles to solve the problem
(B) **Representing and formulating**	**(B1)** Building a shared representation and negotiating the meaning of the problem (common ground)	**(B2)** Identifying and describing tasks to be completed	**(B3)** Describing roles and team organisation (communication protocol/rules of engagement)
(C) **Planning and executing**	**(C1)** Communicating with team members about the actions to be/being performed	**(C2)** Enacting plans	**(C3)** Following rules of engagement, (e.g. prompting other team members to perform their tasks)
(D) **Monitoring and reflecting**	**(D1)** Monitoring and repairing the shared understanding	**(D2)** Monitoring results of actions and evaluating success in solving the problem	**(D3)** Monitoring, providing feedback and adapting the team organisation and roles

Note: The 12 skill cells have been labelled with a letter-number combination referring to the rows and columns for ease of cross-referencing later in the document.

The three major CPS competencies are described below.

1) Establishing and **maintaining shared understanding.** Students must have an ability to identify mutual knowledge (what each other knows about the problem), identify the perspectives of other agents in the collaboration, and establish a shared vision of the problem states and activities (Cannon-Bowers and Salas, 2001; Dillenbourg, 1999; Dillenbourg and Traum, 2006; Fiore and Schooler, 2004). This includes the student's ability to monitor how his or her abilities, knowledge, and perspectives interact with those of the other agents and in relation to the task. Theories of discourse processing have emphasised the importance of establishing a common ground in order for communication to be successfully achieved (Clark, 1996; Clark and Brennan, 1996), so this is also a skill that is essential to CPS. Students must also be able to establish, monitor and maintain the shared understanding throughout the problem-solving task by responding to requests for information, sending important information about tasks completed, establishing or negotiating shared meanings, verifying what each other knows, and taking actions to repair deficits in shared knowledge. These skills involve the student's own self-awareness of proficiencies in performing the task, recognising their own strengths and weaknesses in relationship to the task (metamemory), and recognising the other agents' strengths and weaknesses (transactive memory).

2) Taking appropriate action to solve the problem. Students must be able to identify the type of CPS activities that are needed to solve the problem and to follow the appropriate steps to achieve a solution. This includes efforts to understand the problem constraints, create team goals for the solution, take action on the tasks, and monitor the results in relation to the group and problem goals. These actions may include communication acts, such as explaining, justifying, negotiating, debating and arguing in order for complex information and perspectives to be transferred and for more creative or optimal solutions to be achieved. The constraints and rules of engagement differ for the different types of CPS activities, such as jigsaw problems (where individuals have different knowledge that needs to be pooled; Aronson and Patnoe, 1997), collaborative work (Rosen and Rimor, 2009), and argumentative debates in decision making (Stewart, Setlock and Fussell, 2007). A proficient collaborative problem solver is able to recognise these constraints, follow the relevant rules of engagement, troubleshoot problems and evaluate the success of the problem-solving plan.

3) Establishing and **maintaining group organisation.** A team cannot function effectively without organising the group and adapting the structure to the problem-solving task. Students must be able to understand their own role and the roles of the other agents, based on their knowledge of who is skilled at what in the team (transactive memory), follow the rules of engagement for their role, monitor the group organisation, and facilitate changes needed to handle communication breakdowns, obstacles to the problem and performance optimisation. Some problem situations need a strong leader in the group whereas other problems require a more democratic organisation. A competent student can take steps to ensure that agents are completing tasks and communicating important information. This includes providing feedback and reflecting on the success of the group organisation in solving the problem.

Underlying these three competencies are specific skills that can be individually assessed within collaborative tasks. The assessment is developed ensuring that the skills shown in the 12 cells of the CPS matrix (Table 7.1) are all measured across different tasks. Together the assessment tasks cover the three major competencies and the four component processes.

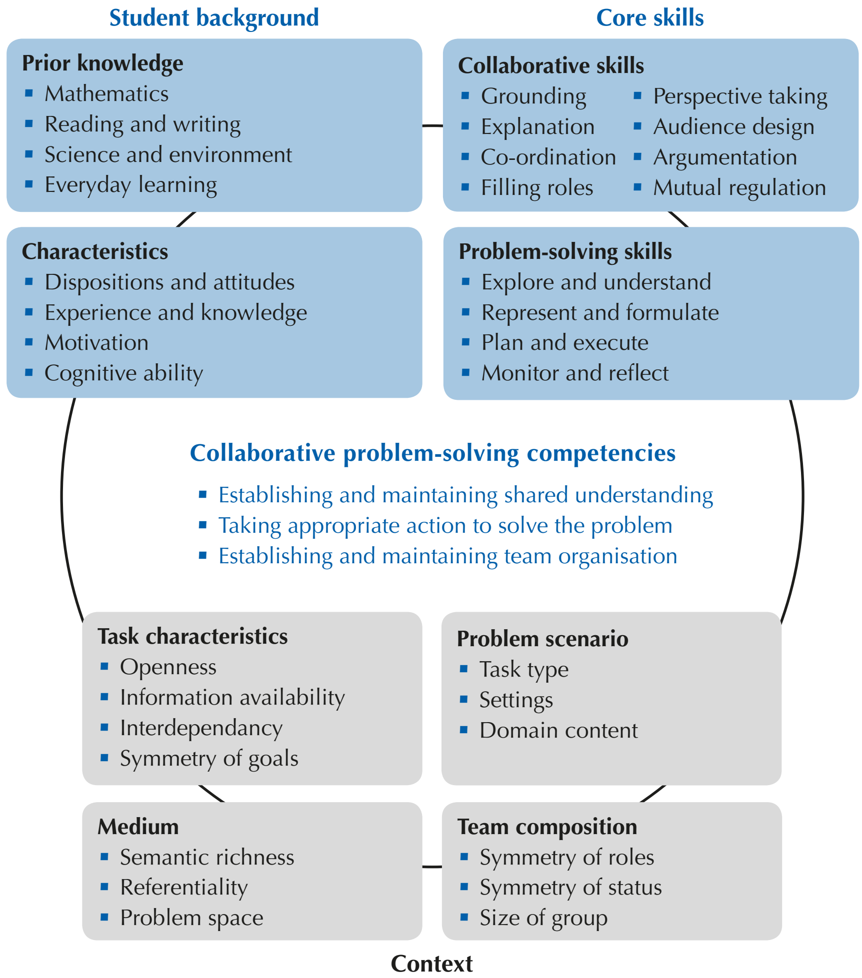
Overview of the domain

Figure 7.1 provides a schema of the salient factors that influence collaborative problem-solving competency, as well as the cognitive and social processes that comprise the skills within collaborative problem-solving contexts, as defined for PISA 2015. The core skills are described above; additional details on the role of student background and task-context factors are provided below.

Student background

A student's prior knowledge and experiences are factors that influence collaboration and problem-solving processes. A student's knowledge of a particular domain, for example of mathematics, the sciences, reading, writing and ICT skills, as well as everyday knowledge, influences the student's capacity to collaborate to solve a problem. Available research indicates that problem-solving strategies rely on domain knowledge to some extent (Funke and Frensch, 2007; Healy et al., 2002; Lee and Pennington, 1993; Mayer, 1992; Mayer and Wittrock, 1996). The assessment uses problem situations and contexts relevant to 15-year-old students that tap generalised problem-solving skills, but do not rely on specialised knowledge. The assessment assumes basic rather than advanced abilities in reading and use of computer interfaces as well as a basic knowledge of science, mathematics and the world. This is similar to the approach adopted in the PISA 2012 individual problem-solving assessment in the selection of problem contexts.

Figure 7.1 ■ **Overview of factors and processes for collaborative problem solving in PISA 2015**



Student characteristics, such as interpersonal skills, attitudes, emotions, personality factors (e.g., the "Big Five" factors of openness, conscientiousness, extraversion, agreeableness and neuroticism) and motivation can all affect individual and collaborative problem-solving success (e.g. Avery Gomez, Wu and Passerini, 2010; Jarvenoja and Jarvela, 2010; Morgeson, Reider and Campion, 2005; O'Neill, Goffin and Gellatly, 2012). Cognitive abilities, such as working-memory capacity, logical reasoning and spatial ability similarly all contribute to CPS. While these core characteristics may influence CPS competence, the PISA 2015 CPS cognitive assessment does not specifically measure factors such as attitude, emotions, motivation or specific domain knowledge. It is, however, intended that the most critical factors are measured as part of the PISA 2015 background questionnaire (see the section "Considerations for contextual questionnaire" below).

The framework assumes that most 15-year-old students have sufficient cognitive and social abilities to complete the CPS tasks. This is a safe assumption from the perspective of research in psychological development. From the standpoint of cognitive and brain development, these students are at an age when most of them are capable of hypothetical reasoning and abstract thought (Bjorklund, forthcoming; Fischer, 1980; Piaget, 1983); from the perspective of social development, they are at an age when most students can consider the perspectives of others and have acquired a large range of socialisation skills (Bjorklund, 1997; Flavell et al., 1968). These capabilities are necessary for being able to establish and maintain a shared understanding in the group, take actions towards a joint goal, and monitor results of collaborative actions.

There is some question as to whether different cultures uniformly value students initiating actions and communications, as opposed to responding to requests and questions. However, taking initiative in appropriate contexts is an important skill at the higher level of CPS competency and therefore relevant to PISA 2015. In the assessment, team members can vary in taking on different task roles, but are not assigned a social status. The assumption is that adopting different roles

in collaborative work and problem solving is acceptable in different cultures; in contrast, social status differences may limit taking initiative in some but not all cultures, thereby imposing a cultural bias. For example, in some cultures there are social customs where it is awkward for an employee to communicate with a boss by asking a question, making a request or evaluating what the boss does. These differences are avoided in the contexts of the PISA 2015 assessment. In contrast, the team members in the problem scenarios have equal status but assume different roles, which is presumed to be acceptable in all cultures and essential to CPS.

Context: Problem scenarios, team composition, task characteristics and medium

The problem scenarios and context in which the problem is solved have a number of psychological dimensions that can affect the type of collaboration and the collaborative competencies employed. These dimensions specify the context of the problem to be solved, the availability of information, the relationships among the group members, and the types of problems.

A meaningful collaborative interaction and motivating experience in assessment does not emerge spontaneously, but requires careful structuring of the collaboration to promote constructive interactions. For example, effective collaboration is characterised by a relatively symmetrical structure with respect to knowledge, status and goals (Dillenbourg, 1999), but the roles and tasks of the different group members may be very different. Symmetry of knowledge occurs when all participants have roughly the same level of knowledge, although they may have different perspectives. Symmetry of status involves collaboration among peers rather than interactions involving facilitator relationships. Finally, symmetry of goals involves common group goals rather than individual goals that may conflict (Rosen and Rimor, 2009).

Assessment items are designed so that successful performance on the task requires collaboration and interdependency among the participants. For example, in many types of problems (i.e., jigsaw, hidden profile [when the information available to the human is not complete at the beginning of the task]), each team member has a piece of information and only together can they solve the problem. These problems are dynamic rather than static because important information accrues during the course of interacting with others. Moreover, problems are designed to provide a graceful degradation of the quality of the solution, so partial or suboptimal solutions receive partial credit. Another example consists of consensus-building tasks, where there are limited resources but a group must bargain and converge on a solution that satisfies the needs of different stakeholders. Information among participants may also conflict, requiring sharing and then resolving the information in order to determine what information best solves the problem (debate).

The assessment items also consider the types of problems that groups of young people must solve, both within a formal school setting and in work contexts in order to be productive in society. A problem scenario provides the situational context in which a problem is applied. For example, within a consensus-building task, a classroom scenario may involve reaching a solution on how to prepare a PowerPoint presentation in a class when students bring different information to the group. Another scenario may be a negotiation task that involves global policies of citizens in a culture, such as a debate on where to build a new school.

The medium of a CPS item defines aspects such as its richness, referentiality and cost of grounding. For example, an item can be graphically rich, providing an immersive environment that simulates a classroom or workplace, or it could be a simple interface providing only a text description of a problem and means to communicate with the group. An item's context may have high referentiality to the outside world and real-world contexts, versus being more abstract, with little reference to external knowledge. An item can have a greater or lesser degree of cost of grounding, depending on how easy it is for members of the group to communicate with each other and find common ground. Finally, an item can have a shared problem space where the actions of each team member are explicitly apparent, such as when working on a shared document; in other scenarios, information about team members' actions might be implicit, for example, when working on separate tasks and reporting back to the group via the communication channel.

The 2012 individual problem-solving framework provides a structure for considerations of aspects of task characteristics, such as ill-defined vs. well-defined, and static vs. dynamic problems. Collaborative problem solving tends to be inherently interactive, interdependent and dynamic (Blech and Funke, 2010, 2005; Klieme, 2004; Wirth and Klieme, 2004). This provides greater challenges to assessment methods as there is much less control over the progress towards solutions, a much wider range of potential problem states, and complexities in tracking problem states. To the extent that any individual in a group depends on other individuals, there is some level of uncertainty in the control over the tasks, making it difficult for most problem types to be fully defined. Thus, a problem may be well-defined from the standpoint of the designer of the problem, but ill-defined at some points from the perspective of one or more group participants. Most or all of the problems also have different phases that can reflect variations in these context dimensions.

Table 7.2 elaborates the schematic representation of Figure 7.1 by providing an overview of the context dimensions and states that can affect the difficulty of the CPS task. In the context of a PISA assessment, it is impossible to assess all of the factors shown in Table 7.2, let alone the large number of combinations of factors; therefore the CPS assessment items constitute a sampling of the total domain by keeping many factors fixed and varying only a few. The framework identifies those factors that are most central to the definition of CPS. More specifically, PISA 2015 CPS concentrates on the collaboration skills to a greater extent than the problem-solving skills needed to solve the particular problem. Consequently, problems vary across low, medium and high difficulty with respect to collaboration skills, while problem-solving skills range from low to medium difficulty.

Table 7.2 Collabrative-problem solving context dimensions

Context	Dimension	States
Problem scenario	Task type	E.g. Jigsaw, consensus building, negotiation
	Settings	Private vs. public Technology vs. non-technology School (formal) vs. non-school (informal)
	Domain content	E.g. Math, science, reading, environment, community, politics
Team composition	Size of group	2 or more (including the student)
	Symmetry of status of team members	Symmetrical vs. asymmetrical
	Symmetry of roles: Range of actions available to each team member	Symmetrical vs. asymmetrical
Task characteristics	Openness (c.f. PISA PS 2012)	Well-defined vs. ill-defined
	Information availability: Does the student receive all necessary information at once? (c.f. PISA PS 2012)	Static vs. dynamic
	Interdependency: Student A cannot solve problem without student B's actions	Low to high
	Symmetry of goals	Group vs. individual
	Distance to solution (from beginning state to goal state)	Small, medium or large
Medium	Semantic richness	Low to high
	Referentiality to the outside world	Low to high
	Communication medium cost of grounding Interdependency: Student A cannot solve problem without student B's actions	Low to high
	Problem space: Does the student get information about other team members' actions?	Explicit vs. implicit

ASSESSING COLLABORATIVE PROBLEM-SOLVING COMPETENCY

There has been substantial research on the development of assessment methods for individual problem solving (the focus of PISA 2012), but work in assessment and training methods for collaborative problem solving is much less developed. As such, there are no established, reliable methods for large-scale assessments of individuals solving problems in a collaborative context and no existing international assessments in wide use. Although ATC21s addresses collaborative problem-solving skills, no measurement at the individual level has yet been reported (Griffin, Care and McGaw, 2011).

Given the overall matrix sampling design used in PISA, where estimates of country-level competency per domain depend on the covariance structure across the domains to be assessed, observations need to address this ability in individuals. Measurements at the individual level can only be obtained if all variables apart from the individual are controlled. Group-level measurements are highly dependent on group composition and the individual skills of the participants (Kreijns, Kirschner and Jochems, 2003; Rosen and Rimor, 2009). Fairly assigning a competency level to individuals working in a group where all group members can vary is impossible, because each individual's display of observable behaviour depends on the behaviour of the other group members.

Further, there are few well-elaborated national or international standards for training or assessing collaborative problem-solving skills. There are, however, a number of research studies, smaller-scale assessments and theoretical work that can inform the development of a reliable large-scale assessment of collaborative problem solving. Annex 7.B provides a deeper review of existing frameworks and assessment approaches.

It has therefore been decided to place each individual student in collaborative problem-solving situations, where the team member(s) with whom the student has to collaborate is/are fully controlled. This is achieved by programming computer agents.

Structure of the assessment

In the PISA 2015 main study, each student is assigned one two-hour test form composed of four 30-minute "clusters". Each form comprises one hour (two clusters) of science, the major domain, with the remaining time assigned to either one or two of the additional domains of reading, mathematics and CPS, according to a rotated test design. Three clusters of material for the CPS assessment were designed and are included in the main study.

CPS units range from 5- to 20-minute collaborative interactions within a particular problem scenario. Multiple measurements of communications, actions, products and responses to probes can be performed within each unit. These measures can be thought of as corresponding to individual items. For example, an item could be a single communication or action taken by a student at a particular point in the problem, the content of a longer sequence of communications and/or actions made by a student, or the correctness of the solution produced. Between 5 and 30 separate measurements are derived from each unit. Each of these individual items provides a score for one or more of the three CPS competencies. Additional details on scoring and weighting of items is provided below. As the CPS assessment is computer-based, the timing information automatically captured during the field trial is used to determine the actual number of items that can be included in each unit and cluster for the main study.

Measurement of collaboration skills

Collaborative problem solving is inherently an interactive, conjoint, dual-strand process that considers how the student reasons about the problem and how the student interacts with others to regulate the social processes and exchange information. These complex processes present a challenge for consistent, accurate and reliable measurement across individuals and across user populations. The complexity of the potential collaborative interactions with the environment increases when there is an attempt to create compelling problem-solving situations in more realistic environments. Computer-based assessment provides an effective means to control the assessment contexts and to collect and analyse student performance. This level of control reduces the complexity in measurement and allows the assessment to be technically implementable. This section describes the focus of what is measured and how computer-based approaches are used.

PISA 2015 CPS is an assessment of individuals in collaborative problem-solving contexts. Because overall analyses for PISA are performed at the student level, the design reflects measuring individual competencies rather than the overall performance of the group process. The PISA 2015 CPS assessment is not designed to measure individuals' cognitive problem-solving skills specifically, but it does do this to the extent that individual problem-solving skills are expressed through collaboration. As such, there is an indirect link to the 2012 individual problem-solving assessment. The 2015 measurement focuses on assessing the cognitive and social processes underlying collaborative problem-solving skills rather than specific domain knowledge.

The process of solving a problem in a collaborative situation in a computer-based assessment generates a complex data set that contains actions made by the team members, communication acts between the group members, and products generated by the individual and the group. Each can be associated with levels of proficiency for each CPS competency. Because the focus is on the individual, assessment items correspond to measures of the student's outputs, whereas outputs from the rest of the group provide contextual information about the state of the problem-solving process.

Prior research and assessments in CPS have used a number of different methods to measure the quality of the problem-solving products (i.e. outcomes) and processes. These methods use varying approaches to assessing actions, communication and products, including measures of the quality of the solutions and objects generated during the collaboration (Avouris, Dimitracopoulou and Komis, 2003), analyses of log files (files to which a computer writes a record of student activities), quality of intermediate results, paths to the solutions (Adejumo, Duimering and Zhong, 2008), team processes and structure of interactions (O'Neil, Chung and Brown, 1997), quality and type of collaborative communication (Cooke et al., 2003, Foltz and Martin, 2008; Graesser, Jeon and Dufty, 2008), and quality of situation judgements (McDaniel et al., 2001). Additional details regarding research on measurement approaches applied to CPS are provided in Annex 7.B.

Individuals working collaboratively on a problem can change the state of a problem by communicating with each other or performing certain actions. For the purpose of the assessment, actions can be defined as any explicit acts made by the individual that change the state of the collaborative problem. These actions include individual acts, such as placing

a puzzle piece, clicking on a button to start a jointly designed machine, moving a cursor on a display that the other participants can see, or editing a joint document. Each action can be mapped to measures of performance as it relates to either success (or failure) of solving the problem or to a skill identified within the framework. For example, placing a puzzle piece incorrectly indicates failure of enacting a plan (cell C2 of the skills matrix, Table 7.1). Sequences of actions provide deeper information about the problem-solving process. For example, the sequence of students' actions in first varying one part of the problem, then verifying the solution and then taking the next appropriate action, can show skills of monitoring results and evaluating success (D2).

While communication is often classified as an individual collaboration skill, the output of communication provides a window into the cognitive and social processes related to all collaborative skills. Students must communicate to collaborate, and the communication stream is captured and analysed to measure the underlying processes. The analysis of the content and structure of communication streams provides measures of the test-taker's ability to share perspectives, establish mutual goals, negotiate with other team members, and take steps to achieve these goals. For example, communication sent by the student indicating what the student sees on a screen provides an indication of building a shared representation (B1). Taking the initiative to ask other agents to manipulate parts of the problem corresponds to following rules of engagement (C3) and enacting plans (C2). Communication acts and sequences of communication acts can be classified to measure the type and quality of skills that are being enacted by the student.

The output or products of the team's problem-solving process provides a third measure of student performance. A product can be based on intermediate and final solutions to the problem-solving process or the output of a "probe item" which checks a student's understanding of a situation in a particular state. These provide a measure of the success that the actions of collaborative problem solving are being enacted properly and that the group is moving the problem state forward appropriately. The products can also be based on "probes" that are placed within the unit to assess a student's cognitive state relative to the skills in the framework. These probes would stop the simulation and ask the student either a constructed-response or multiple-choice question in order to assess knowledge states, shared understanding and the student's understanding of the other group members' skills, abilities and perspectives. The questions range from small tests of the student's state of understanding to situation judgement tasks that require students to put themselves in the context and communicate the state of the problem externally, such as writing an e-mail to a supervisor. Example probes are shown below.

Table 7.3 Example probes

Probe	Skill assessed
What does A know about what is on your screen?	**(A1)** Discovering perspectives/abilities of team members
What information do you need from B?	**(C1)** Communicating with team members about the actions being performed
Why is A not providing information to B?	**(D1)** Monitoring and repairing the shared understanding
What task will B do next?	**(B2)** Identifying and describing tasks to be completed
Who controls the factory inputs?	**(B3)** Describing roles and team organisation
Write an e-mail to your supervisor explaining whether there is consensus of your group on what to do next.	**(B1)** Building a shared representation and negotiating the meaning of the problem **(B2)** Describing tasks to be completed
Write an e-mail to your group explaining what actions the group will need to do to solve the problem.	**(B2)** Identifying and describing tasks to be completed **(C2)** Enacting plans

These explicit probes are one way of assessing students' proficiencies, but much can be inferred from the particular actions and speech acts that do not explicitly probe these knowledge states. For example, if the student does not know whether another group member is aware of what the student has on his or her screen, the student can ask the member a question that targets the uncertainty. Alternatively, another member can perform an action on the screen and observe whether the student comments on an aberration. Physical acts in a shared physical space are acts of communication, just as words and sentences are. Probes can be multiple choice (selected response) or open-ended (constructed response). However, there is no requirement that constructed response be used for such assessments if the skills can be adequately assessed through the actions, communications and products during the collaboration process. Probe items were developed for the two sample CPS units described below, but there are no probe items in the CPS units developed for the PISA 2015 assessment.

To measure performance, all actions, communications, products and response times are logged throughout the problem-solving process. Any action or communication can be thought of as a representation of a particular state of the problem-solving process. Each state of the problem-solving process can also be linked to the specific collaborative skills that need to be assessed, as defined in the framework's CPS skills matrix (Table 7.1). Therefore, items within a unit represent changes in the state performed by the student either through actions, communications or the products resulting from actions or communications.

For example, to assess "establishing and maintaining shared understanding" during the process of "representing and formulating a problem", the state of the problem has pre-determined communication acts related to establishing common ground on tasks (B1). A student initiating a communication act to establish common ground would show that he or she is performing at the highest level in that aspect of collaboration, which would be reflected in the scoring. A student who establishes common ground only after being prompted by the agent would show that he or she is at the proficient level of the skill. Students who send contextually inappropriate communications or who do not communicate any shared understanding would be scored as being below the proficient level.

Pattern-matching technology is used to process the log files and identify the key aspects of performance corresponding to the competencies. This approach permits fully automated partial-credit scoring against each of the skills from the framework. Although there are measures for skills in each cell of the framework, the scores from these skills are combined to create an overall scale for collaborative problem-solving competency.

The student's physical actions, answers to question probes, and acts of communication selected from a menu can be automatically scored. Probes requiring constructed responses, such as short e-mail communications, would require expert-coding. However, because expert-coded responses are assessed off line, the scoring rubric would need to identify the specific skills and context from the framework to be assessed, and would need to measure the quality of the communication and actions.

Conversational agents

The essence of collaborative problem solving is that team members depend on each other. Success in reaching the solution depends on what each team member brings to the collaborative effort. If one of the members in a team has nothing to offer towards solving a problem that requires contributions from all members, the problem will not be solved. Randomly pairing students with other students would therefore lead to an underestimate of the population's problem-solving skills as the weakest member in each pair would determine the probability of success, the quality of the solution, and the efficacy in dealing with the problem.

Research has shown that group composition has a significant effect on performance, in particular the balance of gender (e.g. Bear and Wooley, 2011), ability (e.g. Wildman et al., 2012), personality (e.g. McGivney, Smeaton and Lee, 2008) and what Webb (1995) terms "status characteristics", e.g. family background, popularity, attractiveness and perceived intelligence. In real life, students must be prepared to work effectively within various types of homogenous and heterogeneous groups and with a range of familiar and unfamiliar group members. However, in an assessment situation, if a student is matched with a problematic group, it can have a detrimental effect on the student's performance, and the validity of the assessment is compromised.

Similarly, some students may be highly stimulated when collaborating with one particular student but demotivated when paired with another student. The only way to obtain a full and valid estimate of an individual's collaborative problem-solving skills would therefore be to pair this individual with a number of different team members, each with a different set of characteristics relevant to collaborative problem solving. To ensure fair measurement, each individual student would need to be paired with the same number of other students displaying the same range of characteristics. As PISA is an international study, caution is needed to ensure that the same variability in student characteristics relevant to collaborative problem solving is found in each participating country.

The approach suggested in the previous paragraph is impractical in the context of a large-scale international assessment. Measurement is therefore operationalised using computer-based agents as a means to assess collaborative skills. Students collaborate with computer-based conversational agents representing team members with a range of skills and abilities. This approach allows the high degree of control and standardisation required for measurement. It further permits placing students in a number of collaborative situations and allows measurement within the time constraints of the test administration.

Students are presented with problem scenarios in designated clusters. Each scenario corresponds to an individual assessment unit. The student is asked to respond to the scenario by playing the role of problem solver alongside agents in the given context. CPS skills are measured through a number of items, where each item is linked to a task or phase in the problem-solving process, which can contain several steps.

In each CPS unit, the student works with one or two group members to solve a problem, with the group members programmed as computer agents providing input in much the same manner as fellow students would do. Across different assessment units, agents are programmed to emulate different roles, attitudes and levels of competence in order to vary the CPS situation the student is confronted with. The conversational agents interact with the student's communications and actions in pre-defined ways as the student moves through different states of the problem. Each state defines particular communication acts that can be performed by the agent or that would be expected as input from the student.

As the student progresses through the problem-solving task, the computer monitors the current states of the problem. With each state, the computer provides a changing set of choices of communication acts that a student can use to create a conversation with the agent group member(s). Differing student responses can cause different actions from the agent, both in terms of changes of the state of the simulation (e.g. an agent adding a piece to a puzzle) or conversation (e.g. an agent responding to a request from the student for a piece of information). Similarly, actions performed by the student during the problem solving, such as placing puzzle pieces or moving an object, are also monitored by the computer in order to track progress on the problem-solving process and record the type of student actions relative to the current context of the problem state.

Conversational agents can be manifested in different ways within a computer environment, ranging from simple chat interfaces to full virtual talking heads with full expressiveness. For the purposes of PISA 2015, enhanced menu-based chat interfaces, interactive simulations (e.g. moving cursors in a shared space that team members can all see and respond to) and other web-like applications provide a broad range of conversational contexts and collaborative interaction.

An adequate assessment of a student's CPS skills requires the student to work with multiple types of groups in order to cover the constructs critical for assessment. The computer environment for PISA 2015 is orchestrated so that students interact with different agents, groups and problem constraints to cover the range of aspects defined in the construct. For example, one situation may require a student to supervise the work of agents, where there is an asymmetry in roles. Other tasks may have disagreements between agents and the student, collaboratively orientated agent team members (e.g. initiates ideas, supports and praises other team members), and agent team members with low collaborative orientation (e.g. interrupts, comments negatively about work of others).

When humans collaborate together, it often takes considerable time to make introductions, discuss task properties and assign roles during the initial phases of CPS activities (e.g. "exploring and understanding" and "representing and formulating") and also to monitor and check up on team members during action phases (Marks, Mathieu and Zaccaro, 2001; Wildman et al., 2012; Zaccaro, Marks and DeChurch, 2011). There is also a danger that a group of humans will spend a lot of time pursuing an unproductive path to a solution during the action phase. Within the assessment situation, computer agents allow rigorous control over the collaborative interaction to obtain a sufficient number of assessment events within the test time constraints using strategic dialogue management and rapid immersion in the collaborative context. For example, a "rescue" agent can redirect the group's course of action when too much time has been expended on a poor solution path.

The control of the progression permits the creation of a sufficient number of observations to assess the student's proficiency in each skill specified in the cells of the framework's CPS skills matrix (Table 7.1), particularly within the exacting time constraints of the test administration.

While it is acknowledged that the PISA 2015 assessment does not directly test students working with other students, the agent-based approach permits controlled testing of the skills that are required for collaboration. By targeting these skills under controlled situations, the use of agents provides a sufficiently valid approach to measurement to allow generalisations about the critical collaboration skills. Annex 7.A provides a review of examples of how agent-based environments have been used to assess collaboration, problem solving, tutoring and group learning.

Collaborative problem-solving task types

The assessment includes different types of collaborative problem-solving tasks that elicit different types of interactions and problem-solving behaviours. A typology of the different tasks might segregate (a) group decision-making tasks

(requiring argumentation, debate, negotiation or consensus to arrive at a decision); (b) group co-ordination tasks (including collaborative work or jigsaw hidden profile paradigms where unique information must be shared); and (c) group-production tasks (where a product must be created by a team, including designs for new products or written reports). It is possible to align these categories to either units or items within a unit at different phases, depending on the constraints of item development. For example, consider the following CPS activities:

> **Consensus building:** the group needs to make a decision after considering the views, opinions and arguments of different members. A dominating leader may prevent a sufficient number of perspectives from being shared with the group, so the decision may be non-optimal. The quality of the decision may also be threatened by "group think", swift agreement among members without considering the complexities of the problem.
>
> **Jigsaw problems:** a method to insure interdependence among problem solvers, which is a condition to measure collaboration. Each group member has different information or skills. The group needs to pool the information and recruit each other's skills in order to achieve the group goal. The group goal cannot be achieved by any one member alone. One social loafer who does nothing can jeopardise the achievement of the group goal.
>
> **Negotiations:** group members have different amounts of information and different personal goals. Through negotiation, select information can be passed so that there can be mutual win-win optimisation that satisfies overall group goals.

Additional types of CPS tasks can be appropriate, provided they provide time-constrained collaborative activities requiring ground rules for taking actions, and they establish and maintain both shared understandings and team organisation.

Distribution of units and items

Units serve as the primary context for collaborative problem-solving activities for the assessments. Table 7.2 shows the CPS context dimensions, illustrating a range of potential contexts, problem situations and media that are part of collaborative problem solving. Manipulating all context dimensions would create a large design space of potential CPS assessment activities. To reduce the design space, a set of primary context dimensions have been identified, based on a consensus of experts, that allow for the development of units that assess the major components of CPS skills. This typology of CPS activities uses four dimensions that occur across units (e.g. a unit has only one value on the dimension) and two dimensions in which the value can change within the unit. The typology is as follows:

Across units

- Private vs. public: The context for a problem is private if the scenario is concerned only with the immediate existing problem situation and the group currently solving it – for example, a problem that involves planning a time for a party under the constraints of the participating group members. A public context involves solving a problem in which there is a larger shared context that relates to the external world – for example, a problem that involves the group deciding on the best location to build a school in an under-resourced area.
- Technological vs. non-technological: A technological problem context involves collaboratively working on solving a problem that uses machinery or computer equipment – for example, a problem that involves discovering how something works (e.g. programming an alarm) or uses technology to complete a task (e.g. operating a machine to manufacture the optimal number of shoes). A non-technological context has a referent in the problem that is not technology-related (e.g. planning a party).
- School vs. non-school: A school context involves problems that are typically encountered in a school, while non-school encompasses potential problems that are encountered outside of the school context, e.g. home, work, etc.
- Symmetrical vs. asymmetrical roles: In a problem with symmetrical roles, each group member has the same role in the problem-solving context and all participate equally. In a problem with asymmetrical roles, different roles are assigned to different people, for example, one group member can be assigned to be a scorekeeper while another is assigned the role of controlling a machine.

Within units

- Task type (for example jigsaw, consensus-building, negotiation): As described in the previous section, different types of tasks elicit different types of problem-solving behaviour and interactions among the participants. Within a particular unit, a task type can change. For example, the unit can start with a hidden profile (jigsaw) and then once all the information is shared, it can become a consensus-building task.

- Dynamic vs. static: The 2012 individual problem-solving framework distinguishes between problems that are static (e.g. information disclosed to the problem solver is complete) and those that are dynamic, in which changes in information and states of the problem are beyond the control of the problem solver. In collaborative problem solving, the start of a problem tends to be dynamic, as information about the problem context and other agents is discovered. However, in the middle of a problem, as the student and agents figure out how to execute the actions and understand the roles of the group, the problem may become static. Thus, student performance can be tracked under both static and dynamic problem-solving contexts within units.

Items and weighting for scoring

Each problem scenario (unit) contains multiple tasks. A task, e.g. consensus building, is a particular phase within the scenario, with a beginning and an end. A task consists of a number of turns (exchanges, chats, actions, etc.) between the participants in the team. A finite number of options leading onto different paths are available to the participants after each turn, some of which constitute a step towards solving the problem. The end of a task forms an appropriate point to start the next task. Whenever the participants fail to reach this point, a "rescue" is programmed to ensure that the next task can start.

From a measurement point of view, each task contains one or more items that can be scored. Each item can be coded in two (dichotomous: 0 or 1) or more (polytomous: 0, 1, ..., m) categories, according to item-coding rubrics. The rescue mentioned above ensures that items are independent. The codes reflect the matrix of skills described in Table 7.1 and the proficiencies described later in Table 7.7.

Each item addresses one of the 12 cells in Table 7.1, i.e. the cell that represents the skill that the item aims to assess. The assessment covers all 12 cells, according to weightings discussed below. For example, some items emphasise exploring common ground (A1 in Table 7.1), others require students to clarify roles (B2), others to enact plans (C2), and yet others to reflect on what went wrong in the group (D3). Therefore, each item score contributes to the score for only one cell of the matrix.

The proposed allocation of weights for item development across the 12 skill cells is shown in Table 7.4. Greatest weight is placed on column 1 and then column 3 as these competencies focus specifically on collaborative skills, while column 2 focuses more on problem-solving behaviour within a collaborative context. The overall weighting of the rows is provided as a general guideline. In the PISA 2012 individual problem-solving assessment, it was found to be difficult to distinguish performance between "exploring and understanding" and "representing and formulating" (Greiff et al., 2012). Therefore, the two rows have been combined to provide a joint total weight. Evidence of performance that would fall within either of the two rows would be allocated towards the weight for those combined skills.

Table 7.4 Target weights by target skills

	Establishing and maintaining shared understanding	Taking appropriate action to solve the problem	Establishing and maintaining team organisation	Total
Exploring and understanding				~40%
Representing and formulating				
Planning and executing				~30%
Monitoring and reflecting				~30%
Total	**40-50%**	**20-30%**	**30-35%**	**100%**

Evidence-centred design

In order to measure CPS skills, a systematic measurement methodology is required that can handle the rich data that are collected in the log files of the computer-based assessment. The evidence-centred design (ECD) framework (Mislevy and Haertel, 2006; Mislevy, Steinberg and Almond, 2003) and its computer-based extensions (Clarke-Midura et al., 2011) provides a foundation for developing computer-based performance assessments to measure CPS skills in PISA 2015.

In the OECD framework, assessment is considered a process of reasoning from imperfect evidence using claims and evidence to support the inferences being made about student proficiency. The ECD process includes (a) identifying potential claims about what constitutes student proficiency; (b) identifying evidence (what behaviours/performances elicit skills being assessed, e.g. what students might select, write, do or produce that will constitute evidence for the claims); and (c) identifying the situations (the tasks or items) that give students the optimal opportunity to produce the desired evidence. The purpose is to develop models for schema-based task authoring and for developing protocols for fitting and estimation of psychometric models.

Evidence statements could be used to (a) ground measurement of student performance in observable products elicited by tasks or items; (b) define the distinction between partial and full expressions of the collaborative problem-solving skills; and (c) serve as a basis to develop a wide variety of useful reporting aspects for researchers analysing PISA data, educators, curriculum developers and other interested stakeholders. For example, Table 7.5 lists some design patterns that can guide the development of task-model templates for collaborative problem solving based on an ECD framework.

Table 7.5 Design patterns based on an evidence-centred design framework

Attribute	Description
Rationale	How/why this design pattern provides evidence about focal skill/competency
Focal CPS skill	The primary CPS skill targeted by this design pattern (e.g. establish and maintain shared understanding)
Additional skills	Other skills that may be required by tasks under this design pattern (e.g. explore and understand)
Potential observations	What students actually do, or make, in which they might produce evidence about skills (e.g. students' argumentation in support to agent's claim)
Potential work products	Products a student might produce to demonstrate CPS skills (e.g. correct selection of a hot spot, multiple choice, constructed response)
Characteristic features of tasks	Aspects of assessment situations that are needed to evoke the desired evidence (e.g. student provided with interesting and engaging context or scenario, visible alignment with a specific CPS skill taken from 2015 CPS assessment framework)
Variable features of tasks	Aspects of assessment situations that can be varied in order to shift difficulty or focus (e.g. difficulty of content, scaffolding)

Considerations for computer delivery

The proposed CPS framework with computer agents is compatible with the current capabilities of the PISA 2015 computer platform. The presentation of materials on the computer displays are all conventional media, such as diagrams, figures, tables, simulations (e.g. a shared space that team members can all see and respond to), windows, canned e-mail messages, icons, multiple-choice items, and so on. The student interacts with the agent(s) via a chat window allowing the student to respond through communication menus. With respect to the student inputs, once again, there are conventional interface components, such as mouse clicks, sliders for manipulating quantitative scales, drag-and-drop, cut-and-paste, and typed text input.

All of these standard interactions are supported by the Question and Test Interoperability authoring tool within the computer-based assessment platform TAO and can be automatically scored. These provide a simple means for students to interact with the assessments without requiring specialised knowledge beyond core ICT skills.

One of the salient features of the CPS interface may be an interface for communication between the student and agents. The platform can support communication modes, such as simulated e-mail, web and chat. For example, the interface for a chat communication contains a communication window with lists of alternative messages that can be sent to agents. There are three to five pre-defined alternative speech acts in a communication window that the student can select (via a click), thereby registering an act of communication. These speech acts may be defined according to the described proficiency levels for each cell of the CPS framework matrix. For instance, one act might ask the agent for clarification because the message was ambiguous (failing to detect ambiguities) or another act might ask the agent if he or she performed what he or she was supposed to perform. The fact that there are a limited number of pre-defined message options makes such a communication facility analogous to conventional multiple-choice items in assessments.

Aside from communicating messages, the student can also perform other types of actions on other interfaces, such as verifying in the environment if an action has been performed by another agent or performing an action that another agent failed to perform. Consequently, a sequence of possible message communications, actions and verbal reflections can be collected throughout the process of collaborative problem solving. These are stored in the computer log file. Most messages sent and actions performed can be automatically scored.

Factors affecting item difficulty

A student's overall proficiency in collaborative problem solving can be coded, scored, scaled and measured after defining the specific behaviours to be evaluated for each item and the conditions under which a student demonstrates those behaviours. These behaviours and conditions identify factors from Table 7.2 that determine the difficulty of items for the different collaborative processes. Table 7.6 shows proficient behaviours and conditions under which the behaviours can be manipulated to create item difficulty.

Table 7.6 Relationship between proficient behaviour and determinants of item difficulty

Collaboration processes	Proficient behaviour (summary)	States
(1) Establishing and maintaining shared understanding	▪ Discovers others' abilities - share information about own ability ▪ Discusses the problem - asks questions, responds to others' questions ▪ Communicates during monitoring and resolution of group work	▪ Amount of explicit prior information about others ▪ Size of group ▪ Openness of problem (well-defined/ill-defined) ▪ Having to initiate vs. being prompted to talk
(2) Taking appropriate action to solve the problem	▪ Understands the type of interaction needed, makes sure to know who does what ▪ Describes and discusses tasks and task assignment ▪ Enacts plans together with others and performs the actions of the assigned role ▪ Monitors and evaluates others' work	▪ Interdependency ▪ Intrinsic complexity of problem ▪ Clarity of problem goal ▪ Openness of problem (well-defined/ill-defined) ▪ Distance to solution ▪ Problem space: Explicit or implicit information about group members' actions
(3) Establishing and maintaining team organisation	▪ Acknowledges and enquires about roles ▪ Follows rules of engagement – complies with plan, ensures others do ▪ Monitors team organisation – notices issues, suggests ways to fix them	▪ Symmetry of roles ▪ Problem space: Explicit or implicit information about group members' actions ▪ Co-operativeness of group members

Considerations for contextual questionnaire

Students' characteristics, their prior experiences of CPS and their attitude towards CPS are considered as affective factors towards their performance in the CPS assessment (Figure 7.1). However, general attitudes towards collaborative problem solving are not assessed directly within the cognitive component of the CPS assessment, but in the background questionnaire. In PISA 2012, some student dispositions related to individual problem solving were measured: openness to learning, perseverance and problem-solving strategies. For 2015, an updated set of constructs was developed to incorporate students' experiences and dispositions towards collaboration.

For the 2015 contextual questionnaire, three general constructs were defined as being of interest for psychometric and educational purposes:

Student characteristics: The composition of personality types in collaborative groups has been shown to be an important predictor of performance, particularly extraversion (McGivney, Smeaton and Lee, 2008). Knowing the personality traits of the students and controlling the traits of the agent-partners means that further research can be done to see what effect the "Big Five" personality types (openness, conscientiousness, extraversion, agreeableness and neuroticism) have on performance.

Experiences and **practices:** Collaborative problem solving is not a traditional domain, in that it is not explicitly taught as a school subject; rather, it is embedded as a practice in the classroom. The extent to which students in different PISA participating countries may be familiar with collaboration may differ; therefore it is important to have supporting data on their familiarity with CPS within the following contexts:

- educational: e.g. classroom and assessment experiences
- out-of-school: e.g. home life and hobbies
- technology-specific: e.g. gaming.

Disposition to CPS: The way in which students' perceive CPS and, in particular, their self-efficacy can also affect their performance. Therefore, the following areas are of interest:

- interest in and enjoyment of collaboration
- value of collaboration skills
- self-perception of CPS ability.

Due to logistical and space constraints in the background questionnaire, only some of these constructs are measured. In addition, some information can be gathered through the optional questionnaires, such as the IT, teacher and parent questionnaires.

Reporting proficiency in CPS

A single score summarises students' overall proficiency in collaborative problem solving. To illustrate what the score means, PISA has adopted an approach to reporting survey outcomes that involves the development of learning metrics, which are dimensions of educational progression. Several levels are distinguished and described along these scales in terms of what students typically know and can do.

Table 7.7 Proficiency scale descriptions for collaborative problem solving

Level	What students can typically do
4	At Level 4, students can successfully carry out complicated problem-solving tasks with high collaboration complexity. They are able to solve problems situated in complex problem spaces with multiple constraints, keeping relevant background information in mind. These students maintain an awareness of group dynamics and take actions to ensure that team members act in accordance with their agreed-upon roles. At the same time, they are able to monitor progress towards a solution to the given problem and identify obstacles to be overcome or gaps to be bridged. Level 4 students take initiative and perform actions or make requests to overcome obstacles and resolve disagreements and conflicts. They can balance the collaboration and problem-solving aspects of a presented task, identify efficient pathways to a problem solution, and take actions to solve the presented problem.
3	At Level 3, students can complete tasks with either complex problem-solving requirements or complex collaboration demands. These students can perform multi-step tasks that require the integration of multiple pieces of information, often in complex and dynamic problem spaces. They orchestrate roles within the team and identify information needed by particular team members to solve the problem. Level 3 students can recognise information needed to solve a problem, request it from the appropriate team member, and identify when the provided information is incorrect. When conflicts arise, they can help team members negotiate a solution.
2	At Level 2, students can contribute to a collaborative effort within a problem space of medium difficulty. They can help solve a problem by communicating with team members about the actions to be performed. They can volunteer information not specifically requested by another team member. Level 2 students understand that not all team members have the same information and are able to consider different perspectives. They can help the team establish a shared understanding of the steps required to solve a problem. These students can request additional information required to solve a given problem and solicit agreement or confirmation from team members about the approach to be taken. Students near the top of Level 2 can take the initiative to suggest a logical next step, or propose a new approach, to solve a problem.
1	At Level 1, students can complete tasks with low problem complexity and limited collaboration complexity. They can provide requested information and take actions to enact plans when prompted. Level 1 students can confirm actions or proposals made by others. They tend to focus on their individual role within the group. With support from team members, and working within a simple problem space, these students can contribute to a problem solution.

Four levels of proficiency are identified and described in Table 7.7 in an overall reporting scale for CPS to enable comparisons of student performance between and within participating countries and economies. The descriptions reflect the items that students performing at each level typically can and cannot perform, and the collaborative problem-solving skills associated with these items.

SUMMARY

Collaborative problem solving (CPS) is introduced in PISA for the first time in 2015. The 2015 definition described here builds on the PISA 2012 individual problem-solving assessment, but extends it into the collaborative domain by incorporating the theoretical bases of individual and group cognition. The four processes of the 2012 individual problem-solving framework have been retained and added to the three core competencies identified for collaborative problem solving to produce a matrix of CPS skills. Each of these skills is defined with levels of proficiency that can be measured by the assessment instrument.

The PISA 2015 definition of CPS competency has its origin in the consideration of the types of problems and collaborative interactions that 15-year-old students face in and outside of the classroom, as well as a consideration for their "preparedness for life" in the workplace and in further studies. The ability of each participant in a group to communicate, manage conflict, organise a team, build consensus and manage progress is crucial to its success; the measurement of these skills is at the heart of the three competencies that will form the reporting scales for the assessment.

This framework for 2015 describes and illustrates the collaborative and problem-solving skills that are assessed in PISA 2015, the knowledge and student characteristics that factor into the assessment, and the contexts, team composition and task types that form the basis of the computer-based assessment (Figure 7.1). The framework also explains the rationale for using computer agents to operationalise the measurement of students' collaborative skills. This should enable measurement of the proficiency levels to quantify student performance in CPS.

GLOSSARY OF TERMS

Term	Explanation
Actions	Any explicit acts made by the individual that change the state of the collaborative problem.
Agent	Either a human or a computer-simulated participant in a CPS group.
Cluster	Several units grouped into a 30-minute block for testing.
Consensus-building	A task type where the group needs to make a decision after considering the views, opinions and arguments of different members.
Conversational agent	Computer-based agents representing team members with a range of skills and abilities.
Cost of grounding	How easy it is for members of the group to communicate with each other and find common ground.
ECD (evidence-centred design)	A framework for developing assessments by reasoning from imperfect evidence using claims and evidence to support the inferences being made about student proficiency (Mislevy and Haertel, 2006; Mislevy, Steinberg and Almond, 2003).
Hidden profile task	See jigsaw.
Item	Each problem scenario is divided into different tasks termed "items". Items are a unit of measurement.
Jigsaw	Also known as hidden profile. A task type where each group member has different information or skills. The group needs to pool the information and recruit each other's skills in order to achieve the group goal. The group goal cannot be achieved by any one member alone.
Log file	File to which the computer writes a record of student activities.
Negotiation	A task where group members have different amounts of information and different personal goals. Through negotiation, select information can be passed so that there can be mutual win-win optimisation that satisfies overall group goals.
Openness	The degree to which a problem is "well-defined" (e.g. all the information is at hand for the problem solver) vs. "ill-defined" (e.g. the problem solver must discover or generate new information in order for the problem to be solved).
Probe	A question that stops the problem scenario to assess a student's cognitive state relative to the skills in the framework, e.g. a multiple-choice question to assess knowledge states, shared understanding.
Problem scenario	The problem that the group must solve. Each scenario involves one or more task types and settings. Each unit contains one scenario.

...

Problem space	The space in which the actions are carried out to solve the problem. Can be explicitly or implicitly visible to team members.
Problem state	Any stage within a problem space. Actions or communication can change the state of a problem to another state that is closer or further from the goal.
Product	Outcomes that provide a measure of the success that the actions are being enacted properly and that the group is moving the problem state forward appropriately.
Referentiality	An item's context may have high referentiality to the outside world and real-world contexts or, at the other end of the spectrum, low referentiality with little reference to external knowledge.
Rescue agent	If students reach an impasse or run out of time, a rescue agent will intervene to take students to the beginning of the next item.
Semantic richness	The degree to which the problem provides a rich, elaborated problem context that relates to the external world.
Settings	The context dimensions of the problem scenario, namely: ▪ private vs. public ▪ technology vs. non-technology ▪ school vs. non-school.
Symmetry of roles	The degree to which team members are assigned similar or different roles in a problem scenario.
Symmetry of status	The degree to which the status of team members is the same or different (e.g. peers vs. supervisor-and-subordinate relationships).
Task	A task is a particular phase within the problem scenario consisting of a number of turns between the participants in the team. From a measurement point of view, a task is an item that can be scored.
Task type	The type of collaborative problem-solving task that elicits different types of interactions and problem-solving behaviours. The three types are: consensus-building, jigsaw, negotiations.
Turns	A set of one or more human actions and/or communications in an item.
Unit	Each unit contains one problem scenario and several items.

References

Adejumo, G., R.P. Duimering and **Z. Zhong** (2008), "A balance theory approach to group problem solving", *Social Networks,* Vol. 30/1, Elsevier, Amsterdam, pp. 83-99.

Aronson, E. and **S. Patnoe** (1997)*, The jigsaw classroom: Building cooperation in the classroom* (2nd ed.), Longman, New York.

Avery Gomez, E., D. Wu and **K. Passerini** (2010), "Computer-supported team-based learning: The impact of motivation, enjoyment and team contributions on learning outcomes", *Computers & Education,* Vol. 55/1, pp. 378-390.

Avouris, N., A. Dimitracopoulou and **V. Komis** (2003), "On evaluation of collaborative problem solving: Methodological issues of interaction analysis", *Journal of Computers in Human Behaviour,* Vol. 19/2 pp. 147-167.

Barth, C.M. and **J. Funke** (2010), "Negative affective environments improve complex solving performance", *Cognition and Emotion,* Vol. 24/7, pp. 1259-1268.

Bear, J.B. and **A. Williams Woolley** (2011), "The role of gender in team collaboration and performance", *Interdisciplinary Science Reviews,* Vol. 36, p. 2.

Binkley, M. et al. (2011), "Defining 21st Century Skills", in Griffin, P., B. McGaw and E. Care (eds.), *Assessment and Teaching 21st Century Skills*, Springer, Heidelberg.

Bjorklund, D.F. (forthcoming), "Cognitive development: An overview", to appear in Zelazo, P.D. (ed.), *Oxford Handbook of Developmental Psychology*, Oxford University Press, Oxford, UK.

Bjorklund, D.F. (1997), "The role of immaturity in human development", *Psychological Bulletin*, Vol. 122, pp. 153-169.

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

Blech, C. and **J. Funke** (2005), "Dynamis review: An overview about applications of the Dynamis approach in cognitive psychology", Deutsches Institut fur Erwachsenenbildung, Bonn.

Brannick, M.T. and **C. Prince** (1997), "An overview of team performance measurement", in Brannick, M.T., E. Salas and C. Prince (eds.), *Team Performance Assessment* and *Measurement: Theory, Methods and Applications*, Lawrence Erlbaum Associates, Mahwah, NJ, pp. 3-16.

Campbell, J. (1968), "Individual versus group problem solving in an industrial sample", *Journal of Applied Psychology,* Vol. 52*,* pp. 205-210.

Cannon-Bowers, J.A. and **E. Salas** (2001), "Reflections on shared cognition", *Journal of Organisational Behavior,* Vol. 22, pp.195-202.

Clark, H.H. (1996), *Using language*, Cambridge University Press, Cambridge.

Clark, H.H. and **S.E. Brennan** (1991), "Grounding in communication", in Resnick, L.B., J.M. Levine and S.D. Teasley (eds.), *Perspectives on socially shared cognition*, American Psychological Association, Washington, DC, pp. 127-149.

Clarke-Midura, J. et al. (2011), *Using evidence centered design to develop performance assessments*, Paper presented at the American Educational Research Association Annual Meeting, New Orleans, LA.

Cooke, N.J. et al. (2003), "Measuring team knowledge: A window to the cognitive underpinnings of team performance", *Group Dynamics: Theory, Research and Practice,* Vol. 7, pp. 179-219.

Darling-Hammond, L. (2011), "Policy frameworks for new assessments'", in Griffin, P., B. McGaw and E. Care (eds.), *Assessment and Teaching 21st Century Skills*, Springer, Heidelberg.

Davis, A., D. Fidler and **M. Gorbis** (2011), *Future Work Skills 2020*, Institute for the Future for University of Phoenix Research Institute, www.iftf.org/futureworkskills2020 (accessed 30 September 2012).

Dillenbourg, P. (ed.) (1999), *Collaborative learning: Cognitive* and *computational approaches*, Advances in Learning and Instruction Series, Elsevier Science, Inc, New York, NY.

Dillenbourg, P. and **D. Traum** (2006), "Sharing solutions: Persistence and grounding in multi-modal collaborative problem solving", *The Journal of the Learning Sciences,* Vol. 15*,* pp.121-151.

Fiore, S.M. and **J.W. Schooler** (2004), "Process mapping and shared cognition: Teamwork and the development of shared problem models", in Salas, E. and S.M. Fiore (eds.), *Team cognition: Understanding the factors that drive process and performance*, American Psychological Association, Washington, DC, pp. 133-152.

Fiore, S.M. et al. (2010), "Toward an understanding of macrocognition in teams: Predicting process in complex collaborative contexts", *The Journal of the Human Factors and Ergonomics Society,* Vol. 53*,* pp. 203-224.

Fiore, S.M. et al. (2008), "Processes in complex team problem solving: Parsing and defining the theoretical problem space", in Letsky, M. et al. (eds.), *Macrocognition in Teams: Theories and Methodologies*, Ashgate Publishers, London.

Fischer, K.W. (1980), "A theory of cognitive development: The control and construction of hierarchical skills", *Psychological Review*, Vol. 87/2, pp. 477-531.

Flavell, J.H. et al. (1968), *The Development of Role-Taking and Communication Skills in Children,* Wiley, New York.

Foltz, P.W. and **M.J. Martin** (2008), "Automated communication analysis of teams", in Salas, E., G.F. Goodwin and S. Burke (eds.), *Team Effectiveness in Complex Organisations* and *Systems: Cross-Disciplinary Perspectives and Approaches*, Routledge, New York.

Forrester (2009), Enterprise and SMB Software Survey, North America and Europe, Q42009.

Franklin, S. and **A.C. Graesser** (1996), "Is it an agent or just a program? A taxonomy for autonomous agents", *Proceedings of the Agent Theories, Architectures, and Languages Workshop*, SpringerVerlag, Berlin.

Funke, J. (2010), "Complex problem solving: A case for complex cognition?", *Cognitive Processes*, Vol. 11, pp. 133-142.

Funke, J. (1998), "Computer-based testing and training with scenarios from complex problem-solving research: Advantages and disadvantages", *International Journal of Selection and Assessment*, Vol. 6, pp. 90-96.

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

Glaser, R. and **R. Linn** and **G. Bohrnstedt** (1997), *Assessment in Transition: Monitoring the Nation's Educational Progress*, National Academy of Education.

Graesser, A.C., M. Jeon and **D. Dufty** (2008), "Agent technologies designed to facilitate interactive knowledge construction", *Discourse Processes*, Vol. 45, pp. 298-322.

Greiff, S., S. Wüstenberg and **J. Funke** (2012), "Dynamic Problem Solving: A new measurement perspective", *Applied Psychological Measurement*, Vol. 36/3, pp. 189-213.

Griffin, P., E. Care and **B. McGaw** (2011), "The changing role of education and schools", in Griffin, P., B. McGaw and E. Care (eds.), *Assessment and Teaching of 21st Century Skills*, Springer, Heidelberg.

Hacker, D.J., J. Dunlosky and **A.C. Graesser** (eds.) (2009), *Handbook of Metacognition in Education*, Erlbaum/Taylor & Francis, Mahwah, NJ.

Halpern, D.F. (2003), *Thought* and *Knowledge: An Introduction to Critical Thinking* (4th edition), Lawrence Erlbaum Associates, Mahwah, NJ.

Healy, A.F. et al. (2002), "Optimizing the durability and generalizability of knowledge and skills", in Shohov, S.P. (ed.), *Advances in psychology research*, Vol. 8, Nova Science Publishers, Huntington, New York, pp. 103-174.

Israel Ministry of Education (2011), *Adapting the Educational System to the 21st Century*, national plan, Israel Ministry of Education, Jerusalem.

Jarvenoja, H. and **S. Jarvela** (2010), "Emotion control in collaborative learning situations: Do students regulate emotions evoked by social challenges", *British Journal of Educational Psychology*, Vol. 79/3, pp. 463-481.

Kanter, R.M. (1994), "Collaborative Advantage: The Art of Alliances", *Harvard Business Review*, Vol. 72/4, pp. 96-108.

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

Klein, C., R.E. DeRouin and **E. Salas** (2006), "Uncovering workplace interpersonal skills: A review, Framework, and research agenda", in Hodgkinson, G.P. and J.K. Ford (eds.), *International Review of Industrial and Organizational Psychology*, Vol. 21, Wiley & Sons, Ltd, New York, pp. 80-126.

Kreijns, K., P.A. Kirschner and **W. Jochems** (2003), "Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: A review of the research", *Computers in Human Behavior*, Vol. 19, pp. 335-353.

Lee, A.Y. and **N. Pennington** (1993), "The effect of experience on a cross-domain transfer of diagnostic skill", *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society*, Erlbaum, Hillsdale, NJ.

Marks, M.A., J.E. Mathieu and **S.J. Zaccaro** (2001), "A temporally based framework and taxonomy of team processes", *Academy of Management Review*, Vol. 26, pp. 355-376.

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

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

Mayer, R.E. (1990), "Problem solving", in Eysenck, M.W. (ed.), *The Blackwell Dictionary of Cognitive Psychology*, Blackwell, Oxford, England.

Mayer, R.E. and **M.C. Wittrock** (1996), "Problem solving transfer", in Calfee, R. and R. Berliner (eds.), *Handbook of Educational Psychology*, Macmillan, New York, pp. 47-62.

McDaniel, M.A. et al. (2001), "Use of situational judgment tests to predict job performance: A clarification of the literature", *Journal of Applied Psychology*, Vol. 86, pp. 730-740.

McGivney, S., A.F. Smeaton and **H. Lee** (2008), "The effect of personality on collaborative task performance and interaction", in, Bertino, E. and J.B.D. Joshe (eds.), *Collaborative Computing: Networking, Applications and Worksharing*, Springer.

Ministry of Education Singapore (2008), "Opening address by Dr Ng Eng Hen, Minister for Education and Second Minister for Defence", *International Conference on Teaching and Learning with Technology (iCTLT)*, https://heulab1.wikispaces.com/file/view/ICTLSpeech.pdf (accessed 3 August 2017).

Mislevy, R.J. and **G. Haertel** (2006), *Implications of Evidence-Centered Design for Educational Testing* (Draft PADI Technical Report 17), SRI International, Menlo Park, CA.

Mislevy, R.J., L.S. Steinberg and **R.G. Almond** (2003), "On the structure of educational assessments", *Measurement: Interdisciplinary Research and Perspectives*, Vol. 1, pp. 3-62.

Morgeson, F.P., M.H. Reider and **M.A. Campion** (2005), "Selecting individuals in team settings: The importance of social skills, personality characteristics, and teamwork knowledge", *Personnel Psychology*, Vol. 58, pp. 583-611.

Nash, J. et al. (2003), "Training the transdisciplinary scientist: A general Framework applied to tobacco use behaviour", *Nicotine and Tobacco Research* , Vol. 5, pp. S41-S53.

National Research Council (2011a), *A Framework for K-12 Science Education: Practices, Crosscutting Concepts and Core Ideas*, The National Academies Press, Washington, DC.

National Research Council (2011b), *Assessing 21st Century Skills*, National Academies Press, Washington, DC.

O'Neil, H.F. (1999), "Perspectives on computer-based performance assessment of problem solving: Editor's introduction", *Computers in Human Behavior*, Vol. 15, pp. 255-268.

O'Neil, H.F., S.H. Chuang and **E.L. Baker** (2010), "Computer-based feedback for computer-based collaborative problem-solving", in Ifenthaler, D., P. Pirnay-Dummer and N.M. Seel (eds.), *Computer-Based Diagnostics and Systematic Analysis of Knowledge*, Springer-Verlag, New York, NY.

O'Neil, H.F., S. Chuang and **G.K.W.K. Chung** (2003), "Issues in the computer- based assessment of collaborative problem solving", *Assessment in Education*, Vol. 10, pp. 361-373.

O'Neil, H.F., G. Chung and **R. Brown** (1997), "Use of networked simulations as a context to measure team competencies", in O'Neil, H.F. Jr. (ed.), *Workforce Readiness: Competencies and Assessment*, Lawrence Erlbaum Associates, Mahwah, NJ, pp. 411-452.

O'Neill, T., R. Goffin and **I. Gellatly** (2012), "The knowledge, skill, and ability requirements for teamwork: Revisiting the Teamwork-KSA Test's validity", *International Journal of Selection and Assessment*, Vol. 20/1, pp. 36-52.

OECD (2011), *The OECD Guide to Measuring the Information Society*, http://browse.oecdbookshop.org/oecd/pdfs/free/9311021e.pdf (accessed 12 January 2012).

OECD (2010), *PISA 2012 Field Trial Problem Solving Framework*, www.oecd.org/dataoecd/8/42/46962005.pdf (accessed 29 August 2011).

OECD (2009), *Problem Solving in Technology-Rich Environments: A conceptual Framework*, OECD Education Working Paper No. 36, http://search.oecd.org/officialdocuments/displaydocumentpdf/?doclanguage=en&cote=edu/wkp(2009)15 (accessed 30 September 2012).

OECD (2003), *The PISA 2003 Assessment Framework: Mathematics, Reading, Science and Problem Solving Knowledge* and *Skills*, www.oecd.org/edu/preschoolandschool/programmeforinternationalstudentassessmentpisa/33694881.pdf (accessed 30 September 2012).

Piaget, J. (1983), "Piaget's theory", in Mussen, P. (ed.), *Handbook of Child Psychology* (4th edition), Vol. 1, Wiley, New York.

Roschelle, J. and **S.D. Teasley** (1995), "The construction of shared knowledge in collaborative problemsolving", in O'Malley, C.E. (ed.), *Computer-Supported Collaborative Learning*, SpringerVerlag, Berlin, pp. 69-97.

Rosen, Y. and **R. Rimor** (2012), "Teaching and assessing problem solving in online collaborative environment", in R. Hartshorne, T. Heafner and T. Petty (eds.), *Teacher Education Programs and Online Learning Tools: Innovations in Teacher Preparation*, Information Science Reference, Hershey, PA, pp. 82-97.

Rosen, Y. and **R. Rimor** (2009), "Using collaborative database to enhance students' knowledge construction", *Interdisciplinary Journal of E-Learning* and *Learning Objects*, Vol. 5, pp. 187-195.

Salas, E., N.J. Cooke and **M.A. Rosen** (2008), "On teams, teamwork, and team performance: discoveries and developments", *Human Factors*, Vol. 50, pp. 540-548.

Salas, E. et al. (1992), "Toward an understanding of team performance and training", in Swezey, R.W. and E. Salas (eds.), *Teams: Their Training* and *Performance*, Ablex, Norwood, pp. 3-29; Salas, E. and S.M. Fiore (eds.) (2004), *Team Cognition: Understanding the Factors That Drive Process and Performance*, American Psychological Association, Washington, DC.

Schwartz, D.L. (1995), "The emergence of abstract dyad representations in dyad problem solving", *The Journal of the Learning Sciences*, Vol. 4, pp. 321-354.

Sonnentag, S. and **I. Lange** (2002), "The relationship between high performance and knowledge about how to master cooperative situations", *Applied Cognitive Psychology*, Vol. 16, pp. 491-508.

Stahl, G. (2006), *Group cognition: Computer Support for Building Collaborative Knowledge*, MIT Press, Cambridge, MA.

Stewart, C.O., L.D. Setlock and **S.R. Fussell** (2007), "Conversational argumentation in decision-making: Chinese and U.S. participants in face-to-face and instant-messaging interactions", *Discourse Processes*, Vol. 44, pp. 113-139.

Webb, N.M. (1995), "Group collaboration in assessment: Multiple objectives, processes, and outcomes", *Educational Evaluation and Policy Analysis*, Vol. 17/2, pp. 239-261.

Wildman, J.L. et al. (2012), "Trust development in swift starting action teams: A multilevel framework", *Group & Organization Management*, Vol. 37, pp. 138-170.

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

Wuchty, S., B.F. Jones and **B. Uzzi** (2007), "The Increasing Dominance of Teams in Production of Knowledge", *Science*, Vol. 316/5827, pp.1036-1039, http://dx.doi.org/10.1126/science.1136099.

Zaccaro, S.J., M.A. Marks and **L.A. DeChurch** (eds.) (2011), *Multiteam Systems: An Organization Form for Complex, Dynamic Environments*, Taylor & Francis.

Annex 7.A

STUDIES ON CONVERSATIONAL AGENTS

There is a broad spectrum of computer-based agents that have been used in tasks that involve tutoring, collaborating learning, co-construction of knowledge and collaborative problem solving. (See Table 7.8 for examples of operationally implemented systems). These agents provide a range of techniques that can be potentially incorporated in CPS assessments. At one extreme there are fully embodied conversational agents in a virtual environment, with speech recognition capabilities embedded in a serious game (e.g. the *Tactical Language and Culture System*, Johnson and Valente, 2008). Although this might be motivating to 15-year-old students, this solution would be prohibitively costly and impractical to implement in multiple countries.

A less-expensive solution is animated conversational agents that express themselves with speech, facial expression, gesture, posture and/or other embodied actions. Such systems have been developed and tested in dozens of learning environments during the past two decades, such as *AutoTutor* (Graesser, Jeon and Dufty, 2008; VanLehn et al., 2007), *Betty's Brain* (Biswas et al., 2005), *Operation ARIES* (Millis et al., 2011) and *iSTART* (McNamara et al., 2007a). Although these systems have proven successful in facilitating learning in an impressive body of empirical research, there would be major challenges in technology, costs and cultural variations in language and discourse to implement them in PISA 2015. For example, there are considerable differences among countries in language, speech, communication style, dress, facial demeanour, facial expressions, gesture and so on.

A minimalist approach to assessment using agents provides much of the same control as the more interactive agent approaches, while avoiding some of the above complications. Minimalist agents may consist of printed messages in windows on the computer display, such as e-mail messages, chat facilities, print in bubbles besides icons and documents in various social communication media (Rouet, 2006). Some of these forms of agent-based social communication media have already been implemented in PIAAC (OECD, 2009). There would be no speech generation because of concern of variations among dialects. There might be static visual depictions of the agents who send the messages, which is helpful to mitigate confusion on "who says what" when there are multiple agents playing multiple roles. However, such an approach can minimise the depiction of gender, ethnicity and other visual characteristics of agents that present complications of cultural bias and measurement error.

An important consideration is that it is important for the human to pay attention to the agent when the agent communicates, in a fashion that is analogous to a human who takes the floor when speaking and gets noticed. This can be accomplished with a minimalist agent by a dynamic highlighting of messages and windows through colour, flash and co-ordination of messages with auditory signals (Mayer, 2010).

Computer agents can communicate through a variety of channels. The simplest interface would have the student clicking an alternative on a menu of optional speech acts and for there to be a limited number of options (2 to 7). Other possibilities are open-ended responses that range from typing (or speaking) a single word to articulating sentences and composing lengthier essays. The simplest, but still effective, click interface supports online conditional branching to different system and conversational states, depending on the options the human selects.

Open-ended responses of sentences or essays may be incorporated in the CPS items for later assessment by expert human markers; however, online assessment is still impractical because the advances in computational linguistics (Jurafsky and Martin, 2008) and essay grading (Landauer, Laham and Foltz, 2003; Shermis et al., 2010) are limited or non-existent for some languages. Nevertheless, it would be prudent to collect such open-ended responses for a percentage of assessment items in order to advance research and development of automated language-discourse analyses for future generations. An intermediate solution is semi-structured interfaces, when the system proposes "sentence openers" and then the student completes the sentence (e.g., Soller and Lesgold, 2007). The computer agents can adopt different roles (Baylor and Kim, 2005; Biswas et al., 2005; Millis et al., 2011). For example, the student might take the role of midlevel management and communicate with a supervisor agent and a subordinate agent. The computer agent might be a peer, with equal status to the agent, depending on the way the agent is presented to the subject at the beginning of the text.

The number of computer agents can also vary from only one partner in a dyad, to two agents in a triad, to three or more agents in larger group ensembles. The ensembles of agent configurations are essentially unlimited. Triads (a student and two agents) have advantages because the number of agents is small (minimising confusion in agent roles) but affords interesting complexities in social interaction, such as status differences, agents disagreeing with each other, and agents making comments or taking actions that would make sense to a knowledgeable human (Millis et al., 2011; Wiley and Jensen, 2007). It can also be used to measure social conformity, e.g. whether the student would follow the two agents when they agree on a solution for which the human subject has evidence that it is wrong.

An agent-based approach provides a means to assess individuals' competencies. The proposed minimalist approach to the presence of agents is compatible with the tasks developed for PIAAC (2010) in assessments of problem solving in technology-rich environments. While PIAAC focuses on interaction with technology rather than collaboration, the user interface approach would not be that different. The human would receive e-mail messages from different individuals in addition to working with spreadsheets and web-like searches. Contemporary social communication media (e.g., e-mail, chat, blogs, discussion portals) frequently have messages sent by individuals who cannot be seen and who might not even be known by the recipient of a message (National Research Council, 2011). Teenagers are extensive users of these 21st-century communication media so such interfaces have high ecological validity. Companies also are increasingly adopting mediated natural language communication. Artificial agents are ubiquitous in the modern world and are likely to become even more prevalent in the future.

The following table is a summary of studies with conversational agents in tasks that involve tutoring, collaborative learning, co-construction of knowledge and collaborative problem solving. Innovative assessment systems with agents are being developed at Pearson, Education Testing Service and other assessment organisations (e.g. Forsyth et al., 2012).

Table 7.8 Examples of operationally implemented agent-human based training and assessment systems

Tutor agent and human co-constructing answer to difficult question or solution to problem			
AutoTutor, GuruTutor, Why-Atlas	Physics, biology, computer literacy	Agent helps student articulate answers and solutions through natural language interaction with feedback, hints, prompts for information, corrections and assertions of missing information. Learning gains are the same as human tutors.	Graesser et al. (2004) Olney et al. (2012) VanLehn et al. (2007)
Two agents training humans in skills of reading, writing and speaking			
iSTART	Science texts	Teacher and peer agent train students how to generate self-explanations during reading. Computer interprets natural language and gives feedback. The computer improves comprehension and can accurately identify student paraphrases, relevant elaborations, predictions and other categories of speech acts.	McNamara et al. (2007b) McNamara et al. (2006)
Writing Pal	Argument essays	Teacher and peer agent train students how to write essays by interactively scaffolding different phases of writing. Computer gives feedback on writing quality and scaffolds student's mastery of particular writing components.	McNamara et al. (2012)
Tactical Language and Culture Training System	Language learning	Students learn new languages with multiple agents in socio-cultural contexts. Speech recognition is excellent and students learn. Won the DARPA technological achievement award in 2005 for Tactical Iraqi.	Johnson and Valente (2008)
Tutor, mentor and peer agents collaboratively work with the student on reasoning and problem-solving tasks			
Operation ARIES Operation ARA	Scientific methods and reasoning	Tutor and student peer agents hold trialog conversations with the student on scientific reasoning, finding flaws in research studies and asking questions to critique poor research. There is mixed-initiative dialogue in these interactions. Computer agent helps students learn scientific reasoning and can evaluate the quality of student natural language as well as human experts.	Cai et al. (2011) Millis et al. (2011)
Betty's Brain	Biology, environmental science	Student teaches a student agent how to reason and construct a conceptual graph to understand science well enough to take tests. The human and student collaboratively interact in the inquiry process, with a mentor agent stepping in at appropriate points. This teachable agent system helps students learn the skills of self-regulated learning in addition to deep mental models for problem solving and reasoning.	Biswas et al. (2010) Schwartz et al. (2009)
Crystal Island	Biology	Students interact with agents in a virtual world to explore why a disease evolved. The goal is to build enquiry skills.	Rowe et al. (forthcoming)
River City, ECOMove	Ecology	Agents interact with students in groups on problem solving about hazards in ecological systems.	Ketelhut et al. (2007) Metcalf et al. (2011)
MetaTutor	Biology	Students interact with agents to acquire the skills of self-regulated learning and metacognition in the context of biological systems.	Azevedo et al. (2010)
Coach Mike Ada and Grace	Museums of science	Multiple agents interact with patrons in a science museum.	Lane et al. (2011)
BiLAT	Negotiation	Agents help people learn how to negotiate in a different cultural context.	Kim et al. (forthcoming)

References

Azevedo, R. et al. (2010), "Measuring cognitive and metacognitive regulatory processes used during hypermedia learning: Issues and challenges", *Educational Psychologist,* Vol. 45/4, pp. 210-223.

Baylor, A.L. and **Y. Kim** (2005), "Simulating instructional roles through pedagogical agents", *International Journal of Artificial Intelligence in Education,* Vol. 15*,* pp. 95-115.

Biswas, G. et al. (2010), "Measuring self-regulated learning skills through social interactions in a teachable agent environment", *Research* and *Practice in Technology-Enhanced Learning,* Vol. 5/2, pp. 123-152.

Biswas, G. et al. (2005), "Learning by teaching: A new agent paradigm for educational software", *Applied Artificial Intelligence,* Vol. 19/3, pp. 363-392.

Cai, Z. et al. (2011), "Trialog in ARIES: User input assessment in an intelligent tutoring system", in Chen, W. and S. Li (eds.), *Proceedings of the 3rd IEEE International Conference on Intelligent Computing* and *Intelligent Systems*, IEEE Press, Guangzhou, pp. 429-433.

Forsyth, C. et al. (2012), "Learning gains for core concepts in a serious game on scientific reasoning", in Yacef, K. et al. (eds.), *Proceedings of the 5th International Conference on Educational Data Mining*, International Educational Data Mining Society, Chania, Greece, pp 172-175.

Graesser, A.C., M. Jeon and **D. Dufty** (2008), "Agent technologies designed to facilitate interactive knowledge construction", *Discourse Processes*, Vol. 45, pp. 298-322.

Graesser, A.C. et al. (2004), "AutoTutor: A tutor with dialogue in natural language", *Behavior Research Methods, Instruments, and Computers*, Vol. 36, pp. 180-193.

Johnson, L.W. and **A. Valente** (2008), "Tactical language and culture training systems: Using artificial intelligence to teach foreign languages and cultures", in Goker, M. and K. Haigh (eds.), *Proceedings of the Twentieth Conference on Innovative Applications of Artificial Intelligence*, AAAI Press, Menlo Park, CA, pp. 1632-1639.

Jurafsky, D. and **J.H. Martin** (2008), *Speech* and *Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*, Prentice Hall, Upper Saddle River, NJ.

Ketelhut, D. et al. (2007), "Studying situated learning in a multi-user virtual environment", in Baker, E. et al. (eds.), *Assessment of Problem Solving Using Simulation*s, Erlabaum, Mahwah, pp. 37-58.

Kim, J. et al. (forthcoming), "BiLAT: A game-based environment for practicing negotiation in a cultural context", *International Journal of Artificial Intelligence in Education*.

Landauer, T.K., R.D. Laham and **P.W. Foltz** (2003), "Automated Scoring and Annotation of Essays with the Intelligent Essay Assessor", in Shermis, M. and J. Bernstein (eds.), *Automated Essay Scoring: A Cross-Disciplinary Perspective*, Lawrence Erlbaum Publishers, Mahwah, NJ.

Lane, H.C. et al. (2011), "Intelligent tutoring goes to the museum in the big city: A pedagogical agent for informal science education", in Biswas, G. et al. (eds.), *International Journal of Artificial Intelligence in Education*, Springer, Heidelberg, pp. 155-162.

Mayer, R.E. (2010), *Applying the Science of Learning*, Pearson, Upper Saddle River, NJ.

Metcalf, S.J. et al. (2011), "Ecosystem science learning via multi-user virtual environments", *International Journal of Gaming and Computer-Mediated Simulations*, Vol. 3, pp. 86-90.

Millis, K. et al. (2011), "Operation ARIES!: A serious game for teaching scientific inquiry", in Ma M., A. Oikonomou and L. Jain (eds), *Serious Games* and *Edutainment Applications*, Springer, London.

McNamara, D.S. et al. (eds.) (2012), *Applied Natural Language Processing and Content Analysis: Identification, Investigation, and Resolution*, IGI Global, Hershey, PA.

McNamara, D.S. et al. (2007a), "Evaluating self-explanations in iSTART: comparing word-based and LSA algorithms", in Landauer, T. et al. (eds.), *Handbook of Latent Semantic Analysis*, Erlbaum, Mahwah, NJ, pp. 227-241.

McNamara, D.S. et al. (2007b), " iSTART: A web-based tutor that teaches self-explanation and metacognitive reading strategies", in McNamara, D.S. (ed.), *Reading Comprehension Strategies: Theories, Interventions, and Technologies*, Erlbaum, Mahwah, NJ, pp. 397-421.

McNamara, D.S. et al. (2006), "Improving adolescent students' reading comprehension with iSTART", *Journal of Educational Computing Research*, Vol. 34, pp.147-171.

Millis, K. et al. (forthcoming), "Operation ARIES! A serious game for teaching scientific inquiry", in Ma, M. et al. (eds.), *Serious Games and Edutainment Applications*, Springer-Verlag, London, UK.

National Research Council (2011), *Assessing 21st Century Skills*, National Academies Press, Washington, DC.

OECD (2009), *Problem Solving in Technology-Rich Environments: A Conceptual Framework*, OECD Education Working Paper No. 36, http://search.oecd.org/officialdocuments/displaydocumentpdf/?doclanguage=en&cote=edu/wkp(2009)15 (accessed 30 September 2012).

Olney, A. et al. (2012), Guru: "A computer tutor that models expert human tutors", in. Cerri, S.A. and B. Clancey (eds.), *Proceedings of the 11th International Conference on Intelligent Tutoring Systems (ITS 2012)*, Springer-Verlag, Berlin, pp. 256-261.

Rouet, J-F. (2006), *The Skills of Document Use*, Erlbaum, Mahwah, NJ.

Rowe, J. et al. (forthcoming), "Integrating learning, problem solving, and engagement in narrativecentered learning environments", *International Journal of Artificial Intelligence in Education*.

Schwartz, D.L. et al. (2009), "Interactive metacognition: Monitoring and regulating a teachable agent", in Hacker, D.J., J. Dunlosky and A.C. Graesser (eds.), *Handbook of Metacognition in Education*, Routledge Press, pp. 340-358.

Shermis, M.D. et al. (2010), "Automated Essay Scoring: Writing Assessment and Instruction", in Baker, E., B. McGaw and N.S. Petersen (eds.), *International Encyclopedia of Education* (3rd edition)*,* Vol. 4, Elsevier, pp. 20-26.

Soller, A. and **A. Lesgold** (2007), "Collaborative tools in educational practice", in Hoppe, U., H. Ogata and A. Soller (eds.), *The Role of Technology in CSCL: Studies in Technology Enhanced Collaborative Learning*, Springer, pp. 117-120.

VanLehn, K. et al. (2007), "When are tutorial dialogues more effective than reading?", *Cognitive Science,* Vol. 31, pp. 3-62.

Wiley, J. and **M.S. Jensen** (2007), "When small problem solving groups are effective: What leads to successful interactions?", poster presented at the *Annual Meeting of the Society for Text & Discourse*, Glasgow, Scotland.

Annex 7.B

COLLABORATIVE PROBLEM-SOLVING LITERATURE REVIEW

Collaborative problem solving has been investigated in the social sciences for several decades, resulting in a number of theoretical frameworks, models and paradigms of empirical studies. These contributions span the areas of communication, individual and group problem solving, computer-supported co-operative work, and team assessment. Annex 7.B reviews and outlines research from a number of areas that have implications for the design decisions of the CPS assessment. Many studies have assessed particular components of collaborative problem solving, but few have been validated across diverse populations. Moreover, most studies have focused on business, military contexts or college students (Loughry, Moore and Ohland, 2007; Morgeson, Reider and Campion, 2005; Zhuang et al., 2008). Nevertheless, many of the models, studies and frameworks can apply to the 15-year-old PISA population.

Existing frameworks and models for collaborative skills

A number of existing models and frameworks were reviewed in order to conceptualise the key processes involved in CPS. The conceptualisations of collaborative skills differ in the details across the models, but there are a number of correspondences and some convergence. For example, a number divide out different skills related to collaboration and those related to problem solving. Some of these models formed the basis of the development of definitions of the three core competencies adopted in the PISA 2015 CPS framework, namely:

- establishing and maintaining shared understanding
- taking appropriate action to solve the problem
- establishing and maintaining team organisation.

These three core competencies incorporate major processes taken from theoretical frameworks in the literature cited below. Moreover, they correspond to skills that are important for students entering academic and workplace environments and they adhere to the additional constraint that they can be measured in the PISA 2015 assessment.

The ATC21S framework for collaborative problem solving (Griffin, Care and McGaw, 2011) views CPS as a multi-dimensional skill that includes both social or collaborative skills, and cognitive skills. CPS was conceptualised as having five broad skills.

Social skills include:

- **participation and co-operation:** the ability to participate as a member of a group and contribute knowledge
- **perspective-taking:** the ability to place oneself in another's position, which can lead to adaptation, and modification of communication to take the other's perspective into consideration.
- **social regulation:** such as negotiation and resolution of conflicts or misunderstandings.

Cognitive skills include:

- **task regulation:** the identification of the problem space (its description, goals, needs and resources); clear understanding of the problem space supports the skills of social regulation (being aware that the problem space provides a structure within which learners can locate themselves and each other's needs for knowledge or resources)
- **knowledge building:** where unique contributions of information, skills or resources are combined to contribute to the solution of a problem.

The PIAAC Problem Solving in Technology-Rich Environments Framework (OECD, 2009), incorporates several skills related to CPS. It defines problem solving in technology-rich environments as "using digital technology communication tools and networks to acquire and evaluate information, communicate with others and perform practical tasks". It focuses on "ability to solve problems for personal, work and civic purposes by setting up appropriate goals and plans, accessing and making use of information through computers and computer networks" (OECD, 2009). The skills of communicating with others, and setting goals and plans while solving problems are critical in using digital technologies and are core components of collaboration skills.

The Partnership for 21st-Century Skills' framework (Fadel and Trilling, 2009) presents definitions of communication, collaboration skills and problem solving:

- **Communicate clearly**

  Articulate thoughts and ideas effectively using oral, written and non-verbal communication skills in a variety of forms and contexts.

  Listen effectively to decipher meaning, including knowledge, values, attitudes and intentions.

  Use communication for a range of purposes (e.g. to inform, instruct, motivate and persuade).

  Utilise multiple media and technologies, and know how to judge their effectiveness a priori as well as assess their impact.

  Communicate effectively in diverse environments (including multi-lingual).

- **Collaborate with others**

  Demonstrate ability to work effectively and respectfully with diverse teams.

  Exercise flexibility and willingness to be helpful in making necessary compromises to accomplish a common goal.

  Assume shared responsibility for collaborative work, and value the individual contributions made by each team member.

- **Solve problems**

  Solve different kinds of non-familiar problems in both conventional and innovative ways.

  Identify and ask significant questions that clarify various points of view and lead to better solutions.

Stevens and Campion (1994) provide a five-component model of teamwork that includes the following knowledge, ability and skills:

- **conflict solving:** the ability to recognise and encourage useful conflicts and to employ appropriate conflict resolution strategies when conflicts are not useful
- **collaborative problem solving:** the ability to identify situations requiring group problem solving and decision making
- **communication:** listening skills and a willingness and ability to develop open and supportive communication
- **goal setting and performance management:** setting acceptable and appropriate goals and providing feedback
- **planning and task co-ordination:** the ability to co-ordinate activities with other team members.

Another framework suggested by the Center for Research on Evaluation, Standards and Student Testing (CRESST) consists of six measures (O'Neil et al., 2003, 1997):

- **adaptability:** the group's ability to monitor the source and nature of problems, and provide appropriate feedback
- **co-ordination:** a process by which group resources, activities and responses are organised to ensure success
- **decision making:** the ability to integrate information, use judgement, identify possible alternatives, select the optimal solution, and evaluate the consequences
- **interpersonal:** the ability to improve the quality of team member interactions
- **leadership:** the ability to direct and co-ordinate the activities of the team, assess the performance, assign tasks, plan and organise, and establish a positive atmosphere
- **communication:** efficient information exchange between team members in the agreed manner and by using proper terms, and the ability to clarify and acknowledge.

Zhuang et al. (2008) developed a framework that incorporates some of the considerations of the other frameworks to create five content areas:

- **task-related process skills:** collaborative problem solving, decision making, planning and task co-ordination, strategy formulation, co-ordination, goal setting, performance management
- **co-operation with other team members:** adaptability, interpersonal skills
- **influencing team members through support and encouragement**: confidence building, social support
- **resolution of conflicts or disagreements among team members via negotiation strategies:** conflict solving, communication
- **guidance and mentorship of other team members**: leadership, helping others.

Collazos et al. (2007) suggest five system-based indicators of the success in CPS:

- **use of strategies:** the ability of group members to generate, communicate and consistently use a strategy to jointly solve the problem
- **intra-group co-operation:** the application of collaborative strategies during the process of group work
- **reviewing success criteria:** the degree of involvement of group members in reviewing boundaries, guidelines and roles during the group activity
- **monitoring:** the extent to which the group maintains the chosen strategies to solve the problem, keeping focused on the goals and the success criteria
- **the performance of the group:** how good is the result of collaborative work, total elapsed time while working, and total amount of work done.

Interpersonal skills and the attitudinal, behavioural and cognitive components are also considered critical for performing effectively in collaborative situations. Interpersonal skills have been described as a form of social perception and social cognition involving processes such as attention, and decoding in interpersonal situations. These skills can be likened to a form of social intelligence, involving knowledge of social customs, expectations and problem solving (McDonald et al., 2003). Further, they rest on an "ability to understand" behaviours, cognitions and attitudes of individuals (including oneself) and to translate understanding into appropriate behaviour in social situations (Marlowe, 1986). In a dynamic context, interpersonal skills involve continuous correction of social performance based on

reactions of others during social exchanges (Argyle, 1979). This requires a type of monitoring with feedback loops where one continually adapts behaviours based on verbal and non-verbal cues from others involved in the social exchange. In their review of interpersonal skills, Klein, DeRouin and Salas (2006) synthesised the literature to develop a taxonomy of these skills. They defined interpersonal skills as an umbrella term that refers to "goal-directed behaviours, including communication and relationship-building competencies, employed in interpersonal interaction episodes characterised by complex perceptual and cognitive processes, dynamic verbal and non-verbal interaction exchanges, diverse roles, motivations and expectancies" (p. 81).

Discourse in collaborative problem solving

The theoretical framework for problem solving as a social process was developed by Vygotsky (1986, 1978). According to this theory, personal potential could be realised through a process of interaction with and support from the human environment and from various tools. Interpersonal activity when appropriately implemented could lead to intrapersonal mental development. When trying to solve a problem together through the exchange of ideas, a group of learners constructs shared meanings that the individual would not have attained alone. The shared meaning can only be achieved through communication within the group.

Collaborative problem solving is a co-ordinated joint dynamic process that requires periodic communication between group members (i.e. human or computer agents). The discourse that is communicated among the agents provides both a means for the collaboration to occur as well as a means for measuring the collaborative processes. Communication is a primary means of constructing a shared understanding, as modelled in Common Ground Theory (Clark, 1996; Clark and Brennan, 1991). Clark's theory is widely used within CPS literature as a way of addressing the fact that all agents in a problem solving situation must have some sense of shared knowledge in order to solve a task. Some interpretations of this theory have suggested that the original portrayal of grounding must be extended and adapted to group problem solving because of the complex nature of these interactions (Dillenbourg and Traum, 2006).

In order to apply grounding to problem solving, one major discrepancy exists. In the original theory, conversational partners need only achieve a high enough level of shared understanding necessary to facilitate resulting actions (Clark and Wilkes-Gibbs, 1986). However, Schwartz (1995) suggests that effort is required to acquire new knowledge. Dillenbourg, Traum and Schneider (1996) propose that "optimal collaborative effort" is required of all of the participants in order to achieve adequate learning and performance in a collaborative environment. Some empirical evidence from human interactions in collaborative learning environments suggests that persistence in communication may be more important than a common external representation that facilitates grounding, thus supporting the hypothesis of optimal collaborative effort (Dillenbourg and Traum, 2006).

Researchers of Transactive Memory Theory (Barnier et al., 2008; Theiner, 2010; Theiner and O'Connor, 2010) propose that discourse can allow for an externalised representation of knowledge, leading to the emergence of new information from a group beyond that of any one individual. Fiore and Schooler (2004) adopt a view of macrocognition from this proposition and blended two ideas in order to accommodate group problem solving, namely macrocognition with an application of group communication theory (Chi, Glaser and Rees, 1982; Fiore and Schooler, 2004; Hirokawa, 1980; Orlitzky and Hirokawa, 2001). Specifically, the idea of macrocognition in teams focuses on how people of varying backgrounds and expertise are able to interact with other individuals in a fashion that allows for not only a shared representation but also the formation of new knowledge by applying previously acquired information to new situations.

Group communication theory (as functionally applied to decision-making in problem solving) suggests that the degree to which groups contribute time and effort to completing specific subgoals predicts final performance. The first subgoal is to analyse the problem (Campbell, 1968). The next goal is to define the seriousness of the problem or the reason for solving it, followed by identifying causes, and finally consequences to solutions of the problem. Specific concentration to the negative consequences resulting from solutions may increase a group's effectiveness (Orlitzky and Hirokawa, 2001). The need for communication and achievement of subgoals leads to the conclusion that predicting group performance in problem solving tasks relies heavily on the time spent and quality of the interactions of the group members (Fiore et al., 2010). It is important to place students in an environment that facilitates optimal circumstances for both communicating and reaching a solution.

Considerations for problem-solving environments and tasks

Many collaborative problem-solving studies focus on social dilemmas in which group members must resolve a conflict between personal vs. group benefits. For example, the classic Prisoner's Dilemma consists of a scenario in which multiple people are called in by the police and accused of a crime. By co-operating, an individual may receive the least amount of jail time only if all of the other parties do not co-operate. Rational theory predicts that each person will defect (Hargreaves and Varoufakis, 2004) with deleterious effects. Conversely, real-life experiments show that communication leads to higher co-operation in resolving conflicts within groups during this type of problem-solving task (Balliet, 2010; Sally, 1995).

In contrast to asymmetries in goals, hidden profile tasks create asymmetries in information among participants (Stasser and Titus, 1985). A hidden profile task, or "jigsaw" is one where some information is shared among group members but other important parts of the problem are left unshared. That is, all participants possess some information prior to discussion but other pieces of information are distributed separately to members. To effectively solve the problem, all information must be pooled (Stasser, 1988; Stasser and Titus, 2003).

Technology allows investigators to place humans in orchestrated situations and observe their behaviour and reactions. For example, many technological environments are based on naturalistic decision making (Klein, 2008; Klein et al., 1993; Lipshitz et al., 2001; Zsambok and Klein, 1997) in which each individual has his/her own goals, identity, and expertise which must be aligned in decisions

and action in order to reach the end goal that affects both the individual and the group as a whole. According to Fan, McNeese and Yen (2010), naturalistic decision making focuses on decisions that people make in real life. Ill-structured situations can be created in computer-simulated environments in order to conduct group problem solving research. For example, naturalistic decision making has been examined in a computer-mediated environment in order to discover the beneficial aspects of including artificial agents as collaborators during complex problem solving (Fan, McNeese and Yen, 2010).

Problem solving has also been studied with a focus on goal orientation and achievement rather than decision making, an approach derived from operative intelligence theory (Dörner, 1986). This approach concentrates on the cognitive processes of the group members rather than the results of any given task. Researchers analyse behaviour in complex and dynamic situations that are instantiated in computer-simulated environments, as in the case of the microworlds of Tailorshop (Brehmer and Dörner, 1993) and Microdyn (Funke and Frensch, 2007; Greiff, Wüstenberg and Funke, 2012). Tailorshop creates a scenario in which participants must run a business while maintaining multiple and intertwining goals. Microdyn is an artificial environment that can be altered by allowing systematic variation as group members attempt to manage a complex situation with independent subgoals. Because the goals are independent, multiple scenarios can be presented in succession in order to solve the issue of members achieving only one task (Greiff and Funke, 2009).

Measures of teamwork, taskwork and team cognition

Effective teams engage in both taskwork, i.e. efforts focused on accomplishing the required tasks, and teamwork, i.e. efforts aimed at operating cohesively as a unit (McIntyre and Salas, 1995). There have been a number of techniques developed for assessment of these skills. The approaches have included peer evaluation, behavioural observation scales for experts/instructors, peer review questionnaires and surveys. While none are practical for individual measurement for PISA, these methods inform the taskwork, teamwork and interpersonal skills that are critical to measure in collaborative problem solving. Furthermore, many of these same skills being assessed can be measured in a computer-based collection of collaborative problem-solving data. The logs of the communication and actions performed by the students can be directly related to particular skills and processes used in the scales.

Observation scales

Behavioural observation scales are typically assessed through an instructor or rater observing the team interaction or through peer rating. Taggar and Brown (2001) developed behavioural observation scales that focused on interpersonal skills and self-management skills. These were derived from critical incidents to provide context relevant examples. Each member of the team rated each other team member on items related to the following 13 different dimensions:

1. Reaction to conflict
2. Addresses conflict
3. Averts conflict
4. Synthesis of team's ideas
5. Involving others
6. Effective communication
7. Goal-setting/achievement
8. Team citizenship
9. Commitment to team
10. Focus on task-at-hand
11. Preparation for meetings
12. Providing/reaction to feedback
13. Performance management

A subset of specific behaviours relevant to PISA may be derived from these constructs and be captured in an automated fashion.

Team Dimensional Training was developed in the context of complex decision making tasks for the US Navy. It has been validated in a number of settings with a variety of types of teams (e.g. Smith-Jentsch et al., 2008, 1998). In team dimensional training, behavioural observation is used to rate teamwork process along four dimensions:

- **Information exchange:** addresses "what" is passed "to whom" and is meant to capture those processes foundational to a team's ability to develop and maintain shared situation awareness
- **Communication:** addresses "how" information is delivered
- **Supporting behaviour:** captures how teams compensate for one another in service of achieving team objectives
- **Initiative and leadership:** encompasses guidance and direction provided by team members.

A Likert-type scale is used to make performance ratings for each team member. Ratings are typically provided on a Likert-type scale ranging from 1 to 5 (highly ineffective to highly effective). In Table 7.9, the specific components of team dimensional training are listed.

Table 7.9 Components of team dimensional training

Teamwork dimensions	Component behaviours
Information exchange	▪ Passing relevant information to appropriate teammate at the correct time ▪ Gathering information from all relevant sources ▪ Providing periodic situation updates to summarise big picture
Communication delivery	▪ Using proper terminology ▪ Avoiding excess chatter ▪ Speaking clearly and audibly ▪ Delivering complete standard reports containing data in the appropriate order
Supporting behaviour	▪ Offering, requesting and accepting backup when needed ▪ Noting/correcting errors and accepting correction
Initiative and leadership	▪ Explicitly stating priorities ▪ Providing guidance and suggestions to other team members ▪ Providing direction to other team members

Source: Smith-Jentsch et al. (2008), "Guided team self-correction: Impacts on team mental models, processes and effectiveness".

The Comprehensive Assessment of Team Member Effectiveness instrument is a form of peer evaluation developed from a distillation of numerous team behaviour measurement instruments. It uses "peer evaluations" which have been shown to be a reliable and valid indicator of team process in prior research (e.g. Loughry, Moore and Ohland, 2007; Taggar and Brown, 2001). With this form of assessment, following some interaction experiences, peers rate each other's teamwork behaviours using various scales. For example, the 33-item version of the assessment (Loughry, Moore and Ohland, 2007) has been validated in different team problem-solving and decision-making contexts. The teamwork behaviours in the assessment are categorised along the following five dimensions. With this instrument, peers anonymously rate each other based upon their experience in the team interaction. The Comprehensive Assessment of Team Member Effectiveness relies upon Likert-type scales for rating team members on questions relating to four dimensions:

- Contributing to the team's work
- Interacting with teammates
- Keeping the team on track
- Expecting quality
- Having relevant knowledge, skills and abilities

Measures of team cognition

Problem-solving theory states that mental models can be thought of as an organised understanding or mental representation of knowledge. A team mental model, as an extension of an individual mental model, is an organised understanding or mental representation of knowledge regarding a team's goals, tasks, actions, members and performance. This can be related to either taskwork or teamwork. According to team-cognition theory, effective teams hold multiple compatible mental models (Cannon-Bowers, Salas and Converse, 1993) which support both implicit and explicit co-ordination processes.

Four such models have been elaborated. First is an "equipment model" that captures the shared understanding of the technology and equipment necessary for the team task. Second is the task model that captures the understanding of procedures, task contingencies and strategies of the task. Third is the team-interaction model that captures the understanding of the norms of the team, their responsibilities and their interaction patterns. More specifically, this includes roles, responsibilities, information sources, communication channels and role interdependencies, and is essentially "teammate-generic". Last, the teammate model captures understanding of each other's knowledge, skills, and attitudes, that is, their strengths and weaknesses (Lim and Klein, 2006). This is an assessment of teammates' knowledge, skills, abilities and tendencies, and it is essentially "teammate-specific".

Table 7.10 Accuracy and sharedness of mental models

		Accuracy	
		Low-quality mental model	**High-quality mental model**
Sharedness	**Weak agreement**	Worst performance	Accurate but different (e.g. in situations with differing functional roles the team members may have accurate mental models of their own task but not their teammates)
	Strong agreement	Inaccurate but agreed-upon mental models; they may be able to co-ordinate but it would be down the wrong solution paths (e.g., they will get to an incorrect solution rapidly)	Best co-ordination

Source: Lim and Klein (2006), "Team mental models and team performance: A field study of the effects of team mental model similarity and accuracy".

What is critical for problem solving assessments using shared mental model theory is that we must distinguish between accuracy/quality of the mental model and the sharedness/overlap of the mental model. This is illustrated in Table 7.10.

Items used by Lim and Klein (2006) for pairwise comparisons to assess taskwork and teamwork models:

Taskwork mental model survey items:

- Team members are proficient with their own weapons.
- Team members are proficient with other members' weapons.
- Team members are very good at IA drills.
- Team members have a good understanding of the characteristics of the enemy's weapons.
- Team members conduct routine maintenance of their equipment and weapons in the field.
- Team members are allowed to bring their personal weapon home.
- Team members understand the team's task.
- Team members agree on a strategy to carry out the team task.
- Team members understand other members' tasks.
- Tasks in the team are assigned according to individual member's ability.
- Team members are cross-trained to carry out other members' tasks.
- Team members adhere strictly to the team's SOP.
- Team members understand the battlefield situation.
- The team is highly effective.

Teamwork mental model survey items:

- Team members work well together.
- Team members often disagree with each other on issues faced by the team.
- Team members trust each other.
- Team members communicate openly with each other.
- Team members agree on decisions made in the team.
- Team members accept decisions made by the leader.
- Team members interact with one another outside the camp compound.
- Team members back each other up in carrying out team tasks.
- Team members are similar to each other (e.g. personality, temperament and abilities).
- Team members are aware of other team members' abilities.
- Team members are aware of other team members' personal backgrounds (e.g. family background, hobbies and habits).
- Team members know other team members' family members.
- Team members treat each other as friends.
- The team is highly effective.

Early research on team member surveys (Moreland and Myaskovsky, 2000) analysed team interactions to identify examples of awareness of differentiated member knowledge (specialisation), beliefs about team member reliability on that knowledge (credibility) and the effectiveness in orchestrated knowledge processing (co-ordination). More recently, a large portion of the literature on team member surveys has used surveys of member agreement on expertise surrounding these three particular facets of these surveys (see below). This technique was validated in an important series of studies conducted by Lewis (2003). Lewis examined how assessments of specialisation, credibility and co-ordination could be compared against earlier measures of transactive memory (e.g. verbal protocol analysis, recall measures). The Lewis team member survey scale relies upon Likert-type questions for rating team members.

Items from Lewis's (2003) Transactive Memory System Scale:

- **Specialisation:**

  Each team member has specialised knowledge of some aspect of our project.

  I have knowledge about an aspect of the project that no other team member has.

  Different team members are responsible for expertise in different areas.

  The specialised knowledge of several different team members was needed to complete the project deliverables.

  I know which team members have expertise in specific areas.

- **Credibility:**

  I was comfortable accepting procedural suggestions from other team members.

  I trusted that other members' knowledge about the project was credible.

  I was confident relying on the information that other team members brought to the discussion.

  When other members gave information, I wanted to double-check it for myself. (reversed)

  I did not have much faith in other members' "expertise." (reversed)

- **Co-ordination:**

  Our team worked together in a well-co-ordinated fashion.

  Our team had very few misunderstandings about what to do.

  Our team needed to backtrack and start over a lot. (reversed)

  We accomplished the task smoothly and efficiently.

  There was much confusion about how we would accomplish the task. (reversed)

Small teams do not always require a leader, while large groups always need some form of leadership. Much of the small-team collaborations tasks being assessed within PISA would not require leadership by a single individual. The skills however, remain relevant to the CPS framework, incorporating many of the same competencies. Morgeson, Reider and Campion (2010) developed the measure below to examine leadership in teams. This took a functional approach and outlined what types of behaviours in teams are related to leadership. Although this distinguishes between "action" and "transition" phases in teams and the different functions engaged by teams and their leaders it has items examining both "taskwork" and "teamwork". As such, some variant of this may be warranted. That is, even members of a team who are not leaders can engage in leadership behaviours related to both taskwork and to teamwork.

Morgeson et al.'s Team Leadership Questionnaire (2010) includes the following functions:

- **Transition phase leadership functions:**

  Compose team.

  Define mission.

  Establish expectations and goals.

  Structure and plan.

  Train and develop team.

  Sense-making.

  Provide feedback.

- **Action phase leadership functions**

  Monitor team.

  Manage team boundaries.

  Challenge team.

  Perform team task.

  Solve problems.

  Provide resources.

  Encourage team self-management.

  Support social climate.

References

Argyle, M. (1979), "New developments in the analysis of social skills", in Wolfgang, A. (ed.), *Non-Verbal Behavior: Application and Cultural Implications*, Academic Press, New York, pp. 139-158.

Balliet, D. (2010), "Communication and cooperation in social dilemmas: A meta-analytic review", *Journal of Conflict Resolution,* Vol. 54/1*,* pp. 39-57.

Barnier, A.J. et al. (2008), "A conceptual and empirical framework for the social distribution of cognition: The case of memory", *Cognitive Systems Research (Special Issue, 'Perspectives on Social Cognition')*, Vol. 9/1, pp. 33-51.

Brehmer, B. and **D. Dörner** (1993), "Experiments with computer-simulated microworlds: Escaping both the narrow straits of the laboratory and the deep blue sea of the field study", *Computers in Human Behavior,* Vol. 9*,* pp. 171-184.

Campbell, J. (1968), "Individual versus group problem solving in an industrial sample", *Journal of Applied Psychology,* Vol. 52*,* pp. 205-210.

Cannon-Bowers, J.A., E. Salas and **S.A. Converse** (1993), "Shared mental models in expert team decision making", in Castellan, N.J. Jr. (ed.), *Individual and Group Decision Making: Current Issues*, LEA, Hillsdale, NJ, pp. 221-246.

Chi, M.T.H., R. Glaser and **E. Rees** (1982), "Expertise in problem solving", in R.J. Sternberg (ed.), *Advances in the Psychology of Human Intelligence*, Erlbaum, Hillsdale, NJ, pp. 7-75.

Clark, H.H. (1996), *Using Language*, Cambridge University Press, Cambridge.

Clark, H.H. and **S.E. Brennan** (1991), "Grounding in communication", in Resnick, L.B., J.M. Levine and S.D. Teasley (eds.), *Perspectives on Socially Shared Cognition*, American Psychological Association, Washington, DC, pp. 127-149.

Clark, H.H. and **D. Wilkes-Gibbs** (1986), "Referring as a collaborative process", *Cognition,* Vol. 22*,* pp.1-39.

Collazos, C.A. et al. (2007), "Evaluating Collaborative Learning Processes using System-based Measurement", *Educational Technology & Society,* Vol. 10/3, pp. 257-274.

Dillenbourg, P. and **D. Traum** (2006), "Sharing solutions: Persistence and grounding in multi-modal collaborative problem solving", *The Journal of the Learning Sciences,* Vol. 15*,* pp.121-151.

Dillenbourg, P., D. Traum and **D. Schneider** (1996), "Grounding in multi-modal task-oriented collaboration", in Brna, P., A. Paiva and J. Self (eds.), *Proceedings of the European Conference on Artificial Intelligence in Education*, Lisbon, Portugal, pp. 401-407.

Dörner, D. (1986), "Diagnostik der operativen intelligenz", *Diagnostica,* Vol. 32*,* pp. 290-308.

Fadel, C. and **B. Trilling** (2009), *21st Century Skills: Learning for Life in Our Times*, CA: Jossey-Bass, San Francisco.

Fan, X., M. McNeese and **J. Yen** (2010), "NDM-Based cognitive agents for supporting decision-making teams", *Human-Computer Interaction,* Vol. 25*,* pp. 195-234.

Fiore, S.M. and **J.W. Schooler** (2004), "Process mapping and shared cognition: Teamwork and the development of shared problem models", in Salas, E. and S.M. Fiore (eds.), *Team Cognition: Understanding the Factors That Drive Process and Performance*, American Psychological Association, Washington, DC, pp. 133-152.

Fiore, S.M. et al. (2010), "Toward an understanding of macrocognition in teams: Predicting process in complex collaborative contexts", *The Journal of the Human Factors* and *Ergonomics Society,* Vol. 53*,* pp. 203-224.

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

Greiff, S., S. Wüstenberg and **J. Funke** (2012), "Dynamic Problem Solving: A new measurement perspective", *Applied Psychological Measurement*, Vol. 36/3, pp. 189-213.

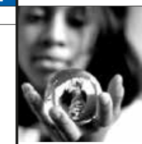
Greiff, S. and **J. Funke** (2009), "On the way to competence levels in dynamic microsystems: The MicroDYN Approach", in Funke, J., J. Wirth and S. Greiff (eds.), *Symposium on Problem Solving, Assessment of Problem Solving Competencies*, paper presented at the EARLI in Amsterdam, The Netherlands.

Griffin, P., E. Care and **B. McGaw** (2011), "The changing role of education and schools", in Griffin, P., B. McGaw and E. Care (eds.), *Assessment* and *Teaching of 21st Century Skills*, Springer, Heidelberg.

Hargreaves, H., P. Shaun and **Y. Varoufakis** (2004), *Game Theory: A Critical Text* (2nd edition), Routledge, New York.

Hirokawa, R.Y. (1980), "A comparative analysis of communication patterns within effective and ineffective decision making groups", *Communication Monographs*, Vol. 47, pp. 312-321.

Klein, G. (2008), "Naturalistic decision making", *Human Factors*, Vol. 50*,* pp. 456-460.



Klein, C., R.E. DeRouin and **E. Salas** (2006), "Uncovering workplace interpersonal skills: A review, framework, and research agenda", in Hodgkinson, G.P. and J.K. Ford (eds.), *International Review of Industrial and Organizational Psychology*, Vol. 21, Wiley & Sons, Ltd, New York, pp. 80-126.

Klein, G. et al. (1993), *Decision making in action: Models* and *methods*, Ablex Publishing, Norwood, NJ.

Lewis, K. (2003), "Measuring transactive memory systems in the field: Scale development and validation", *Journal of Applied Psychology*, Vol. 88, pp. 587-604.

Lim, B.C. and **K.J. Klein** (2006), "Team mental models and team performance: A field study of the effects of team mental model similarity and accuracy", *Journal of Organisational Behavior*, Vol. 27, pp. 403-418.

Lipshitz, R. et al. (2001), "Taking stock of naturalistic decision making", *Journal of Behavioral Decision Making*, Vol. 14, pp. 331-352.

Loughry, M., D. Moore and **M. Ohland** (2007), "Development of a theory-based assessment of team member effectiveness", *Educational* and *Psychological Measurement*, Vol. 67/3, pp. 505-524.

McDonald, S. et al. (2003), "TASIT: A new clinical tool for assessing social perception after traumatic brain injury", *The Journal of Head Trauma Rehabilitation*, Vol. 18/3, pp. 219-238.

McIntyre, R.M. and **E. Salas** (1995), "Measuring and managing for team performance: Emerging principles from complex environments", in Guzzo, R. and E. Salas (eds.), *Team Effectiveness and Decision Making in Organisations*, Jossey-Bass, San Francisco, pp. 149-203.

Marlowe, J.H.A. (1986), "Social intelligence: evidence for multidimensionality and construct independence", *Journal of Educational Psychology*, Vol. 78/1, pp. 52-58.

Moreland R.L. and **L. Myaskovsky** (2000), "Exploring the performance benefits of group training: Transactive memory or improved communication?", *Organisational Behavior and Human Decision Processes*, Vol. 82/1, pp. 117-133.

Morgeson, F.P., D.S. DeRue and **E.P. Karam** (2010), "Leadership in teams: A functional approach to understanding leadership structures and processes", *Journal of Management*, Vol. 36, pp. 5-39.

Morgeson, F.P., M.H. Reider and **M.A. Campion** (2005), "Selecting individuals in team settings: The importance of social skills, personality characteristics, and teamwork knowledge", *Personnel Psychology*, Vol. 58, pp. 583-611.

OECD (2009), *Problem Solving in Technology-Rich Environments: A Conceptual Framework*, OECD Education Working Paper No. 36, http://search.oecd.org/officialdocuments/displaydocumentpdf/?doclanguage=en&cote=edu/wkp(2009)15 (accessed 30 September 2012).

O'Neil, H.F., S. Chuang and **G.K.W.K. Chung** (2003), "Issues in the computer- based assessment of collaborative problem solving", *Assessment in Education*, Vol. 10, pp. 361-373.

O'Neil, H.F., G. Chung and **R. Brown** (1997), "Use of networked simulations as a context to measure team competencies", in O'Neil, H.F. Jr. (ed.), *Workforce Readiness: Competencies and Assessment*, Lawrence Erlbaum Associates, Mahwah, NJ, pp. 411-452.

Orlitzky, M. and **R.Y. Hirokawa** (2001), "To err is human, to correct for it divine: A meta-analysis of research testing the functional theory of group decision-making effectiveness", *Small Group Research*, Vol. 32, pp. 313-341.

Sally, D. (1995), "Conversation and cooperation in social dilemmas: a meta-analysis of experiments from 1958 to 1992", *Rationality & Society*, Vol. 7, pp. 58-92.

Schwartz, D.L. (1995), "The emergence of abstract dyad representations in dyad problem solving", *The Journal of the Learning Sciences*, Vol. 4, pp. 321-354.

Smith-Jentsch, K.A. et al. (2008), "Guided team self-correction: Impacts on team mental models, processes and effectiveness", *Small Group Research*, Vol. 39, pp. 303-327.

Smith-Jentsch, K.A. et al. (1998), "Team dimensional training: A strategy for guided team selfcorrection", in Cannon-Bowers, J.A. and E. Salas (eds.), *Making decisions under stress*, American Psychological Association, Washington, DC, pp. 271-297.

Stasser, G. (1988), "Computer simulation as a research tool: The DISCUSS model of group decision making", *Journal of Experimental Social Psychology*, Vol. 24, pp. 393-422.

Stasser, G. and **W. Titus** (2003), "Hidden profiles: A brief history", *Psychological Inquiry*, Vol. 14/3 and Vol. 14/4, pp. 304-313.

Stasser, G. and **W. Titus** (1985), "Pooling of unshared information in group decision making: Biased information sampling during discussion", *Journal of Personality* and *Social Psychology*, Vol. 48, pp. 1467-1478.

Stevens, M.J. and **M.A. Campion** (1994), "The knowledge, skills and ability requirements for teamwork: Implications for human resources management", *Journal of Management*, Vol. 20/2, pp. 502-528.

Taggar, S. and **T.C. Brown** (2001), "Problem-solving team behaviors: Development and validation of BOS and a hierarchical factor structure", *Small Group Research*, Vol. 32, pp. 698-726.

Theiner, G. (2010), "Making sense of group cognition: The curious case of transactive memory systems", in Christensen, W., E. Schier and J. Sutton (eds.), *ASCS09: Proceedings of the 9th conference of the Australasian society for cognitive science*, Macquarie Centre for Cognitive Science, Sydney, pp. 334-342.

Theiner, G. and **T. O'Connor** (2010), "The emergence of group cognition", in Corradini, A. and T. O'Connor (eds.), *Emergence in Science and Philosophy*, Routledge, New York.

Vygotsky, L.S. (1986), *Thought* and *language*, The MIT Press, Cambridge, MA.

Vygotsky, L.S. (1978), *Mind* and *Society: The Development of Higher Mental Processes*, Harvard University Press, Cambridge, MA.

Zhuang, X. et al. (2008), "Development and validity evidence supporting a teamwork and collaboration assessment for high school students", *ETS Research Report*, 08-50/2, pp. i-51, http://dx.doi.org/10.1002/j.2333-8504.2008.tb02136.x.

Zsambok, C. and **G. Klein** (1997), *Naturalistic Decision Making*, Lawrence Erlbaum Associates, Mahwah NJ.

Annex 7.C

PISA 2015 COLLABORATIVE PROBLEM-SOLVING SAMPLE UNITS

Purpose and scope of sample units

Two collaborative problem-solving (CPS) units were developed as preliminary samples to illustrate the concepts of the assessment framework and show how it might be operationalised. These samples were tried out with a small number of students representing the target testing population in the context of cognitive lab interviews. This confirmed that students could demonstrate the targeted skills when answering the items and that those skills could, therefore, potentially be measured. The samples are not intended to be complete units: they do not cover all item types available, and they do not demonstrate the computer platform used in PISA 2015. These samples will be replaced by released items that contain more detailed information about scoring and student performance.

Both units contain several items, showing how the different competencies in the CPS skills matrix (Table 7.1) are measured. The following assessment and educational principles guided the development of the sample units:

- evidence-centred design (ECD)
- engaging CPS scenarios relevant for 15-year-old students
- phrase chat to operationalise communication between the student and the computer agent (canned words and phrases, appropriate for each situation, are presented in a menu format; the student constructs the dialogue by selecting phrases)
- progression through each unit based on a mapping of the phrase chat and actions possible for each situation; this functionality allows for a standardised CPS assessment of each student
- consideration of cognitive load, colour contrast, and navigation complexity
- scaffolding (embedded "rescue agent" functionality is provided by the computer agent[s] to allow sufficient control over interaction to assure assessment of the full range of CPS proficiencies in the skills matrix)
- clear stimulus material and brief task statements to reduce the dependency on reading proficiency.

To illustrate appropriate coverage of the major CPS skills, one of the units is characterised by a symmetrical nature of collaboration (THE AQUARIUM), while in the second unit the student is assigned as leader of a team with two agents to achieve a common goal (CLASS LOGO). The assessment scenarios include simulations of disagreements between the agent and the student, collaboratively-orientated agent behaviours (e.g. initiating ideas, building consensus, and supporting and praising other team members), and low collaborative agent behaviours (e.g. interrupting other members of the team or commenting negatively about work of others). This allows for a range of situations and team compositions to be presented to the student and thus provides a sufficient dataset for CPS assessment.

Sample CPS unit: THE AQUARIUM

Unit classifications

Context: in-school | outside school
Contents: consensus building, win-win negotiation, **hidden profile (jigsaw) task**
Type of CPS task: decision making | co-ordination | production
Number of agents: two agents, including the student
Target unit timing: 5 minutes | 10 minutes | **15 minutes** | 20 minutes

Unit overview (team composition, problem context and overview of tasks)

In this unit the test-taker and Abby (a computer agent) collaborate to find the optimal conditions for fish living in an aquarium. The test-taker controls three variables (water, scenery and lighting) and Abby controls three other variables (food, fish population and temperature). Within each unit, there are several tasks, each of which may contain one or more assessment items. Scores are accumulated based on the test-taker's performance on individual items.

The first task involves an initial consensus-building discussion between the test-taker and Abby on how to solve the problem (exploring and understanding). Then the team proceeds to a series of collaborative hidden-profile tasks to find the optimal conditions for the fish (representing and formulating, and planning and executing). In the final task, the test-taker monitors and reflects on the collaborative work. The test-taker is told that the number of attempts to solve the problem (known as "trials") is limited to five. The first attempt is set up so that the test-taker will not be able to solve it right away, i.e. the underlying principle of the task forces the test-taker to be involved in at least two trials to gather sufficient data in order to measure CPS skills.

Agent overview

Abby represents collaboratively orientated agent behaviour (e.g. she initiates ideas, builds consensus, and responds to, supports and praises the test-taker). However, in some situations Abby shows misunderstanding of the results and suggests misleading strategies to solve the problem. As long as the test-taker clarifies or repairs misunderstandings or points out the advantages or disadvantages of different strategies, Abby is persuaded. However, if the test-taker does not clarify misinterpretations of results or provide evidence that counters a suggested strategy, Abby will press for a rationale for accepting the strategy.

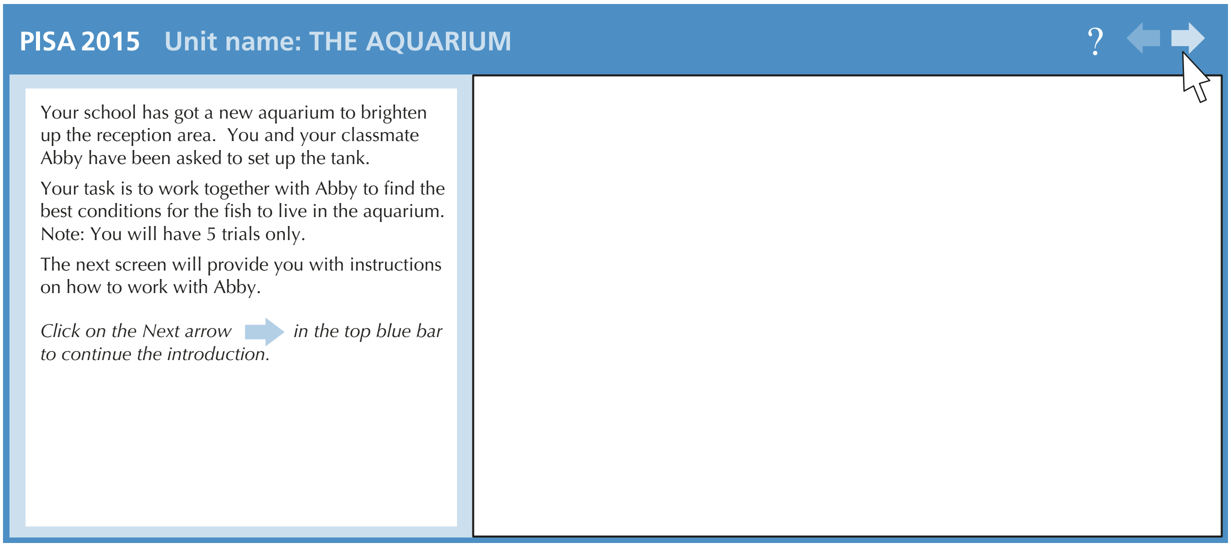
CPS skills

In this unit the test-taker demonstrates CPS skills by establishing a shared understanding of the problem, clarifying misunderstandings and building consensus with a team member on the actions to be performed. The specific cells addressed in the framework matrix from Table 7.1 are described below.

Introduction and orientation

The unit starts with a briefing on the scenario outline and training on the Chat, Control Panel and Task Space areas of the screen. This section is not timed or scored.

Figure 7.2 [1/2] ■ **Sample unit THE AQUARIUM: Introduction**



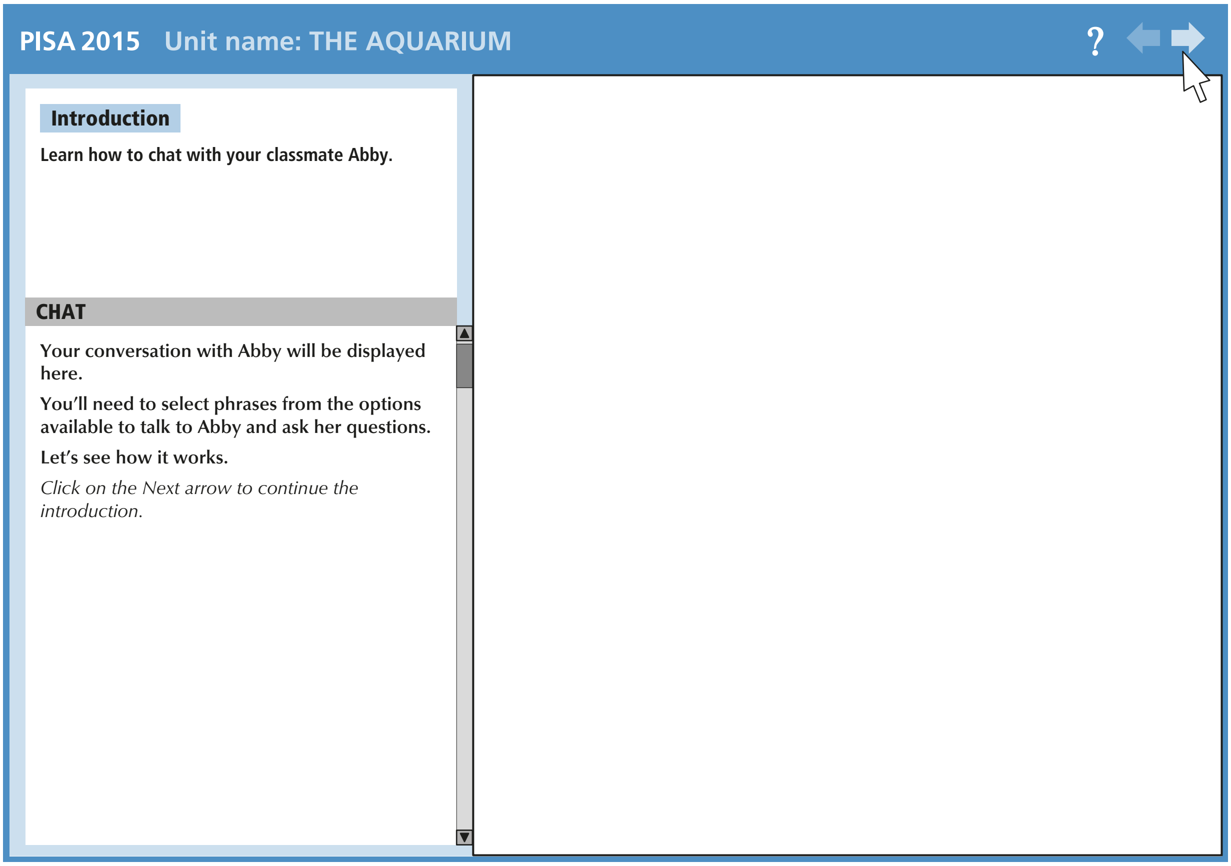
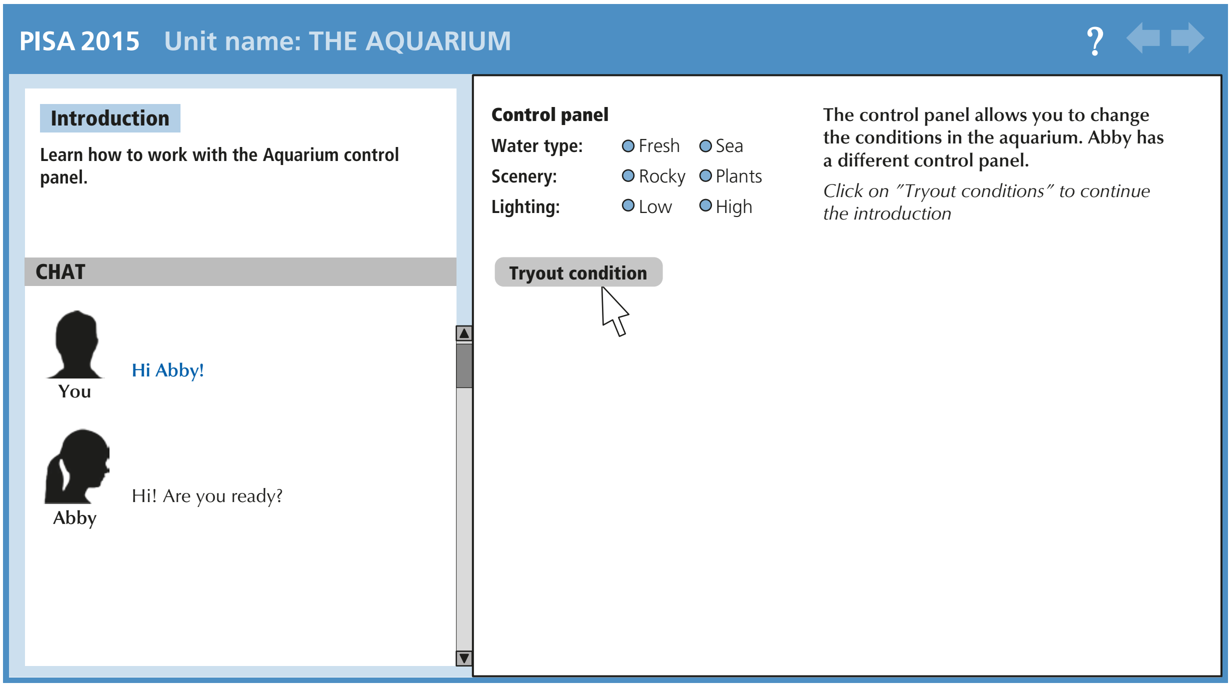
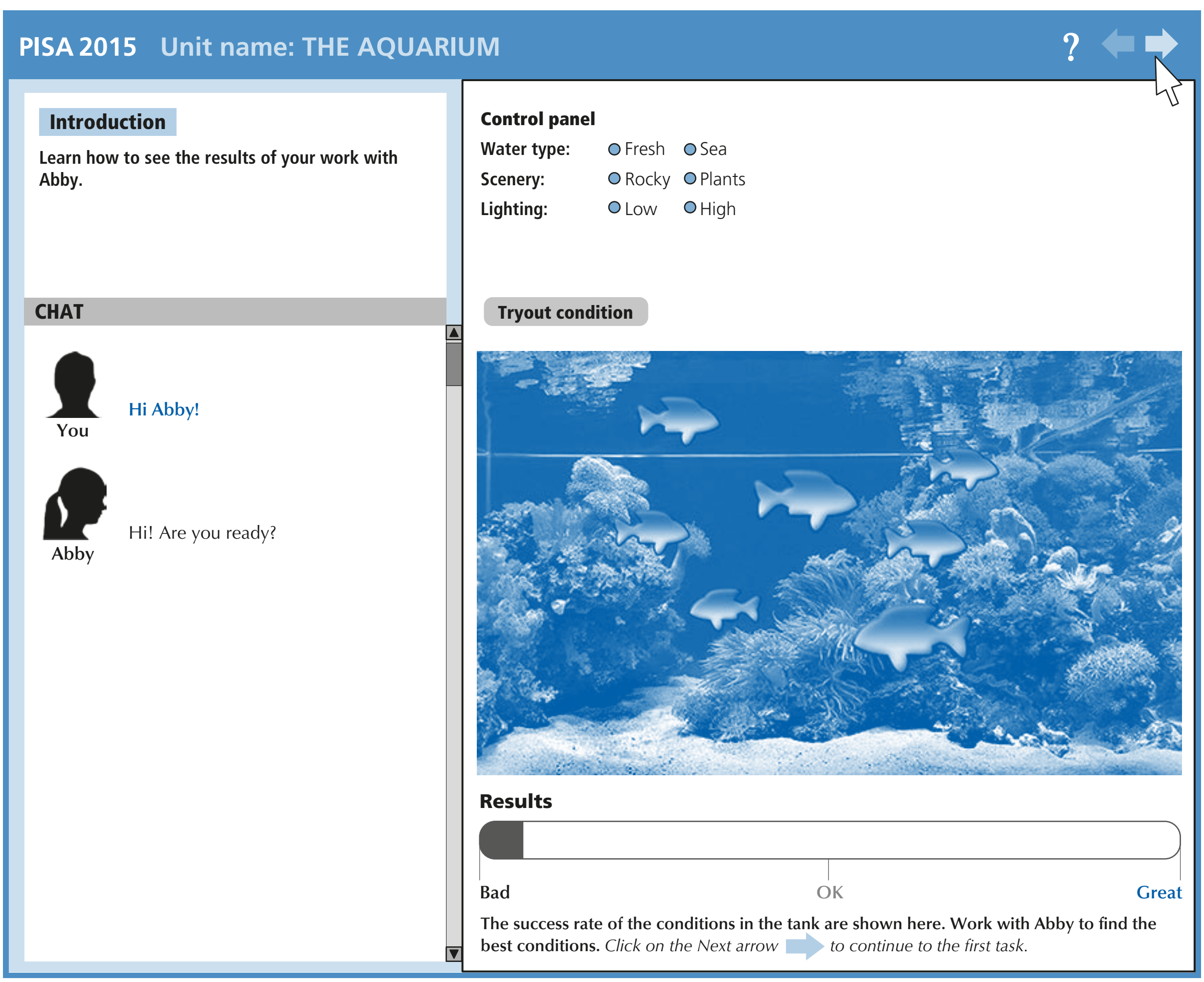


Figure 7.2 [2/2] ■ **Sample unit THE AQUARIUM: Introduction**





Outline of unit tasks

Task 1: Establish shared understanding

Activity

- **Item 1:** The test-taker has to find out what Abby's controls are by asking her. If the test-taker asks, Abby shares her screen (and the test-taker receives one score point for the skill). If the test-taker doesn't ask and tries to move too quickly to actions, then Abby will perform a rescue and offer to share her screen (and the test-taker receives zero score points for the skill).
- **Item 2:** The test-taker has to click on the "share screen" button to reciprocate and allow Abby to see his or her controls. If the test-taker doesn't perform the action within a certain amount of time, then Abby will prompt again.
- **Item 3:** The test-taker offers a plan of how to reach the optimum solution and asks Abby for her point of view. If the test-taker doesn't offer an idea, then Abby prompts. If still no idea is offered, Abby will suggest an idea herself.
- **Item 4:** The test-taker has to ensure that Abby is in agreement (i.e. monitors shared understanding) before clicking on "Next" to try out the new conditions for fish. If the test-taker doesn't offer to click "Next", then Abby will rescue and request or encourage the test-taker to do something. When the test-taker clicks "Next", a pop-up asks if both team members are ready to start the next task. If the test-taker did not agree with Abby beforehand, then Abby can interject here and the test-taker can repair before clicking "Yes" to proceed.

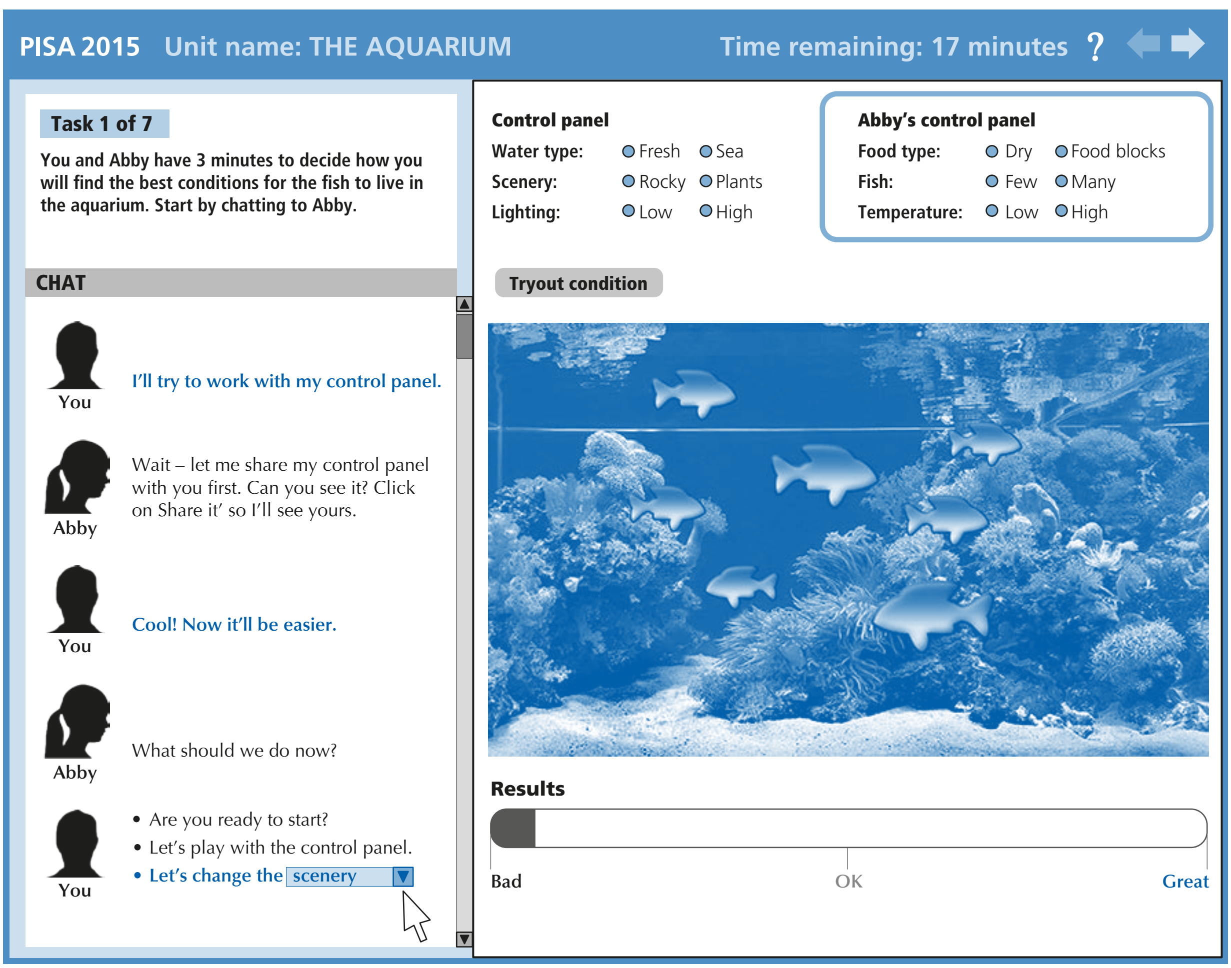
Convergence

The test-taker can see Abby's controls and vice versa. The test-taker and the agent have decided on a plan.

CPS skill(s) assessed across the items within the task:

(A1) Discovering perspectives and abilities of team members; **(A2)** discovering the type of collaborative interaction to solve the problem, along with goals; **(C1)** communicating with team members about the actions to be/being performed; **(B1)** building a shared representation and negotiating the meaning of the problem (common ground).

Figure 7.3 ▪ **Sample unit THE AQUARIUM: Task 1**



Task 2: Enacting plans and monitoring the results

Activity

- **Item 1:** The test-taker monitors whether Abby followed the plan as discussed, while Abby's controls show that she didn't follow the plan. The test-taker shares his or her understanding of the result (fish conditions).
- **Item 2:** The test-taker has to offer a plan of how to proceed (e.g. "let's change this variable"). If the test-taker doesn't offer an idea, then Abby can prompt. If still no idea is offered, then Abby will suggest an idea herself.
- **Item 3:** The test-taker asks Abby for her point of view before implementing the plan. If the test-taker doesn't ask, then Abby shares her view with the test-taker.

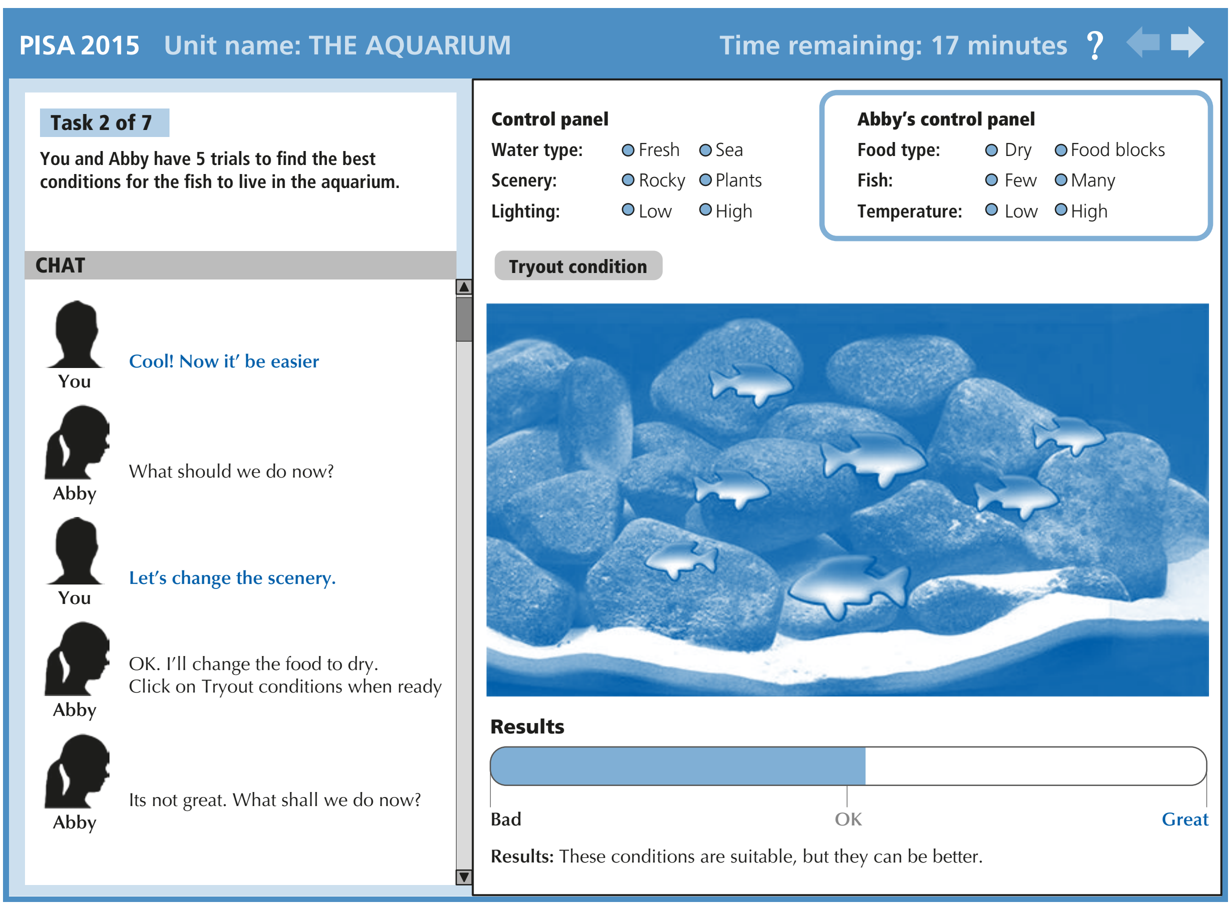
Convergence

There is a change in the aquarium variables. The results of the trial are presented.

CPS skill(s) assessed:

(A1) Discovering perspectives and abilities of team members; **(A2)** discovering the type of collaborative interaction to solve the problem, along with goals; **(C1)** communicating with team members about the actions to be/being performed; **(B1)** building a shared representation and negotiating the meaning of the problem (common ground).

Figure 7.4 ▪ **Sample unit THE AQUARIUM: Task 2**



Task 3: Monitoring and repairing the shared understanding

Activity

- **Item 1:** The test-taker implements the plan as discussed with Abby. The test-taker monitors whether Abby followed the plan as discussed. Abby's controls show that she is following the plan.
- **Item 2:** The test-taker shares his or her understanding of the result (fish conditions).
- **Item 3:** The test-taker repairs Abby's misunderstanding of the result.

- **Item 4:** The test-taker has to offer a plan of how to proceed (e.g. "let's change this variable to start"). If the test-taker doesn't offer an idea, then Abby can prompt. If still no idea is offered, then Abby will suggest an idea herself.
- **Item 5:** The test-taker asks Abby for her point of view before implementing the plan. If the test-taker doesn't ask, then Abby shares her perspective with the test-taker.

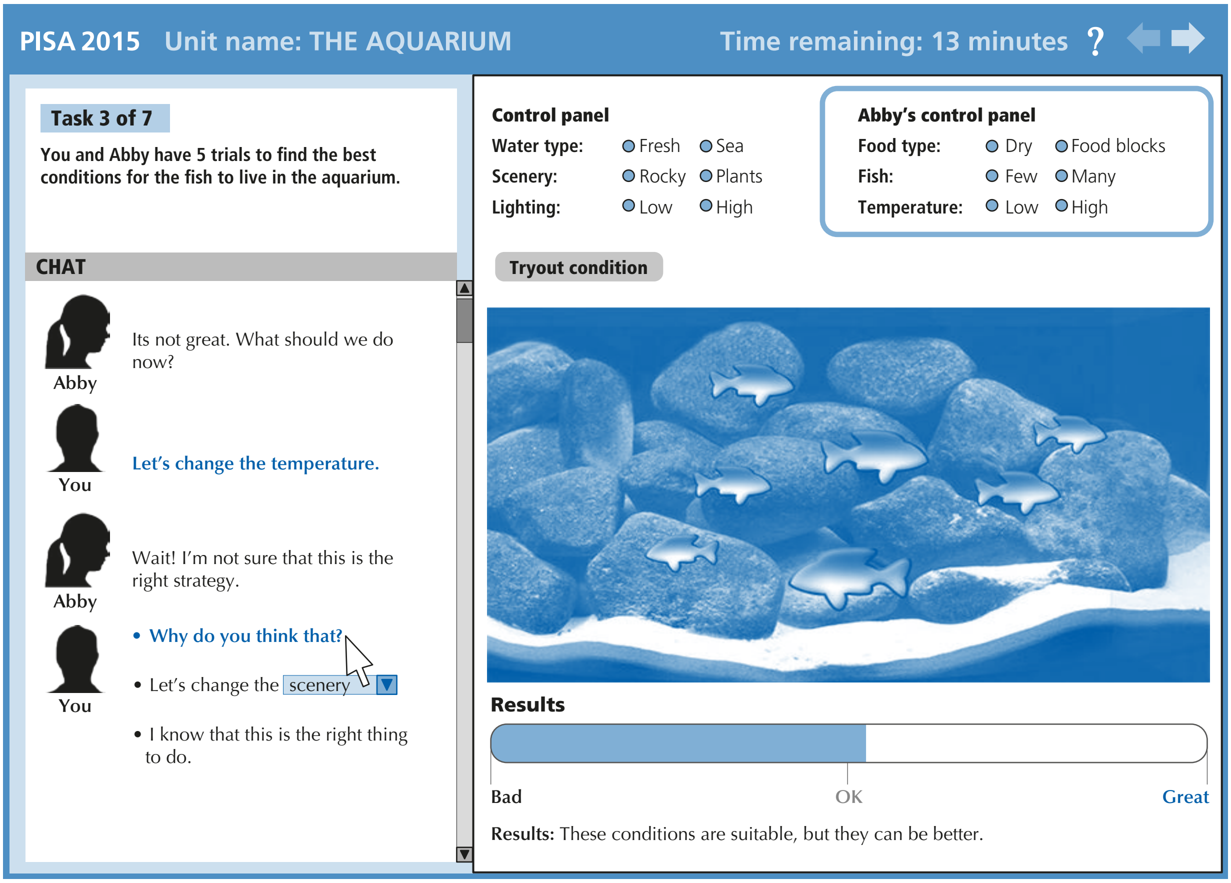
Convergence

There is a change in the aquarium variables. The results of the trial are presented.

CPS skill(s) assessed across the items within the task:

(C2) Enacting plans; (D2) monitoring results of actions and evaluating success in solving the problem; (D1) monitoring and repairing the shared understanding; (C1) communicating with team members about the actions to be/being performed; (B1) building a shared representation and negotiating the meaning of the problem (common ground).

Figure 7.5 ■ **Sample unit THE AQUARIUM: Task 3**



Tasks 4-6

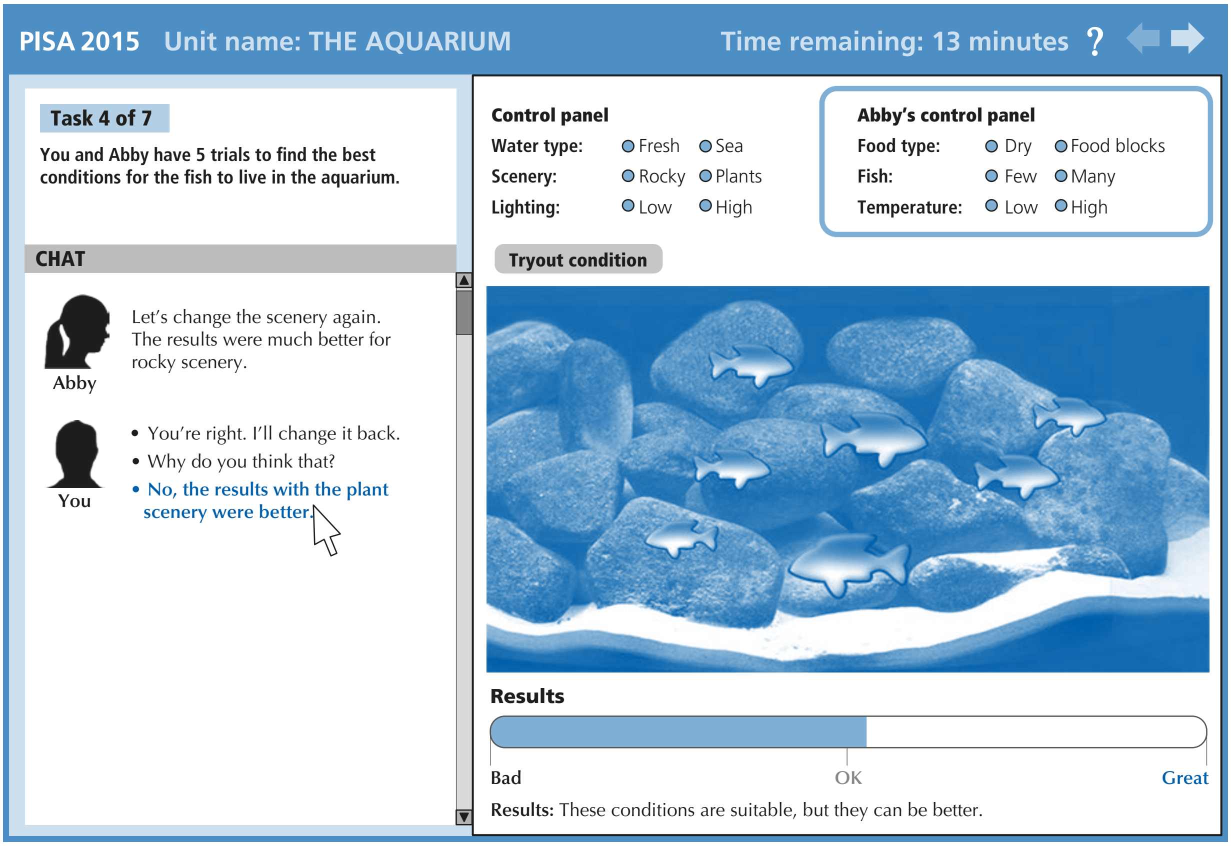
These are only presented if applicable, depending on the test-taker's performance.

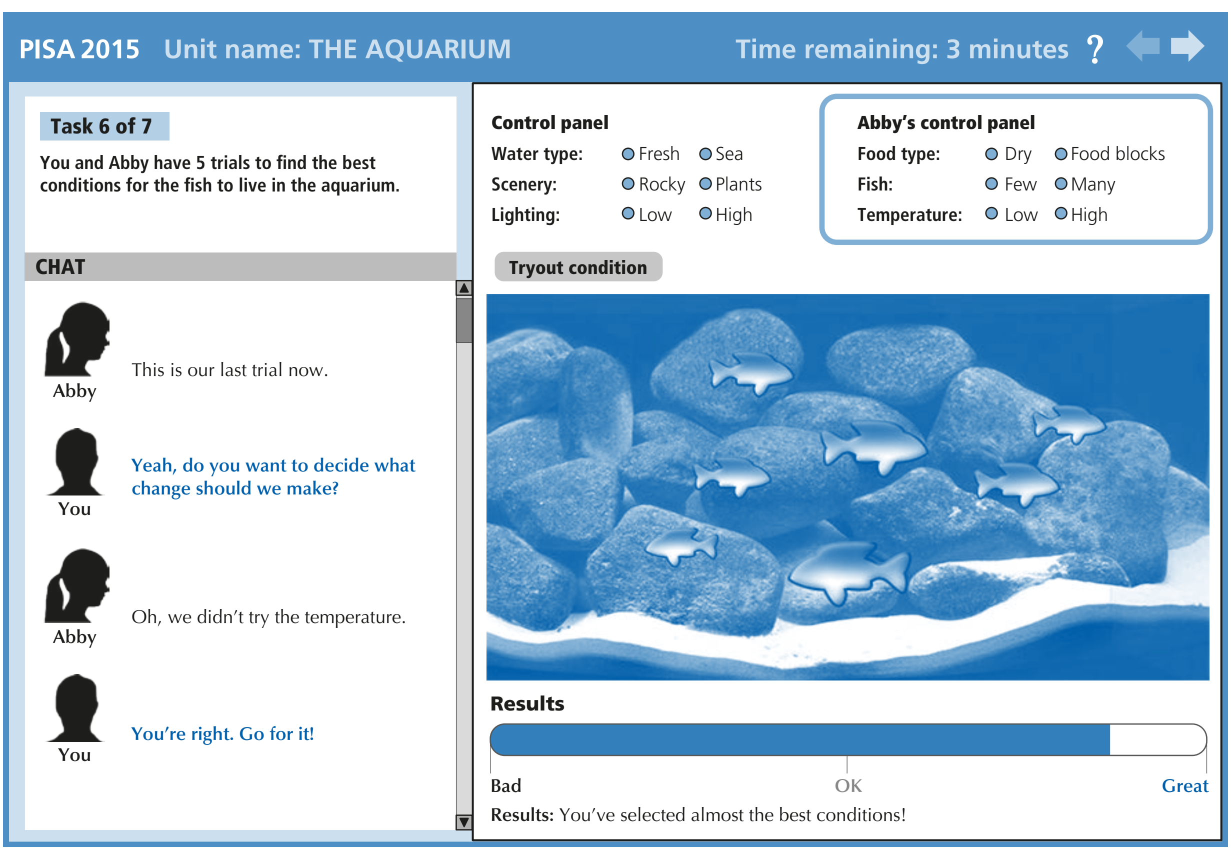
Activity

Optimising the strategy to solve the problem

- **Item 1:** The test-taker implements the plan as discussed with Abby. The test-taker monitors whether Abby followed the plan as discussed. Abby's controls show that she is following the plan.
- **Item 2:** The test-taker shares his or her understanding of the result (fish conditions).
- **Item 3:** The test-taker has to offer a plan of how to proceed (e.g. "let's change this variable"). If the test-taker doesn't offer an idea, then Abby can prompt. If still no idea is offered, then Abby will suggest an idea herself.
- **Item 4:** The test-taker asks Abby for her point of view before implementing the plan. If the test-taker doesn't ask, Abby shares her perspective with the test-taker.

Figure 7.6 ▪ **Sample unit THE AQUARIUM: Tasks 4-6**





<u>Convergence</u>

There is a change in the aquarium variables. The results of the trials are presented.

CPS skill(s) assessed across the items within the task:

(C2) Enacting plans; **(D2)** monitoring results of actions and evaluating success in solving the problem; **(C1)** communicating with team members about the actions to be/ being performed.

As test-takers may make multiple attempts to optimise the strategy to solve the problem, they would receive scores based on the number of attempts, with fewer attempts resulting in higher scores (0-2) for C2. Test-takers would also receive the maximum score achieved across attempts for skills D2 and C1.

Task 7: Providing feedback

<u>Activity</u>

- **Item 1:** The test-taker provides reflective feedback on his or her work with Abby. The test-taker is required to suggest a more collaborative method to promote co-operation with Abby on the task (e.g. talk more to Abby).

<u>Convergence</u>

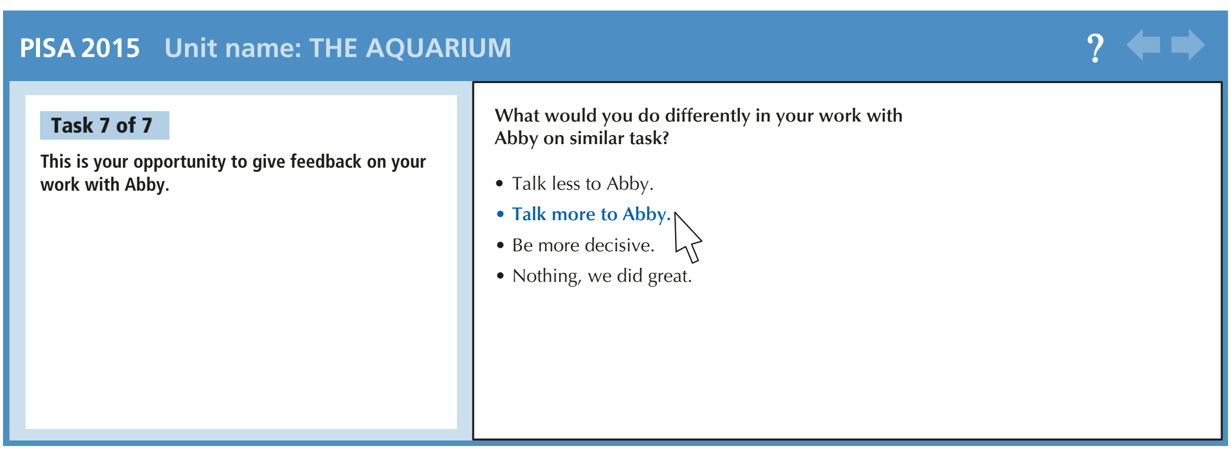
Abby and the test-taker give feedback on the collaborative work.

CPS skill(s) assessed across the items within the task:

(D3) Monitoring, providing feedback and adapting the team organisation and roles.

The question is presented in a multiple-choice format. There is a single optimal answer, which receives full credit. Some of the other options would receive partial credit, and some options would receive no credit.

Figure 7.7 ▪ **Sample unit THE AQUARIUM: Task 7**



Unit measurement profile

At the end of each task, there is a convergence point. This ensures that all students start from the same point and have the same opportunity to score.

Table 7.11 Profile of assessment items within sample unit THE AQUARIUM

Task no.	Item no.	Item short description	Target collaborative problem-solving skill	Data type	Score range (0-x)
1	1	Test-taker finds out what Abby's controls are by asking her.	**(A1)** Discovering perspectives and abilities of team members	Communication	0-2
1	2	Test-taker clicks on "share screen" button to reciprocate and allow Abby to see his or her controls.	**(A2)** Discovering the type of collaborative interaction to solve the problem, along with goals	Action	0-2
1	3	Test-taker offers a plan of how to reach the optimum solution.	**(C1)** Communicating with team members about the actions to be/ being performed	Communication	0-2
1	4	Test-taker asks Abby for her point of view before implementing the plan.	**(B1)** Building a shared representation and negotiating the meaning of the problem (common ground)	Communication	0-2
2	1	Test-taker implements the plan as discussed with Abby.	**(C2)** Enacting plans	Action	0-2
2	2	Test-taker monitors if Abby followed the plan as discussed.	**(D1)** Monitoring and repairing the shared understanding	Communication	0-2
2	3	Test-taker shares his or her understanding of the result (fish conditions).	**(D2)** Monitoring results of actions and evaluating success in solving the problem	Communication	0-2
2	4	Test-taker offers a plan of how to reach the optimum solution.	**(C1)** Communicating with team members about the actions to be/ being performed	Communication	0-2
2	5	Test-taker asks Abby for her point of view before implementing the plan.	**(B1)** Building a shared representation and negotiating the meaning of the problem (common ground)	Communication	0-2
3	1	Test-taker implements the plan as discussed with Abby.	**(C2)** Enacting plans	Action	0-2
3	2	Test-taker shares his or her understanding of the result (fish conditions).	**(D2)** Monitoring results of actions and evaluating success in solving the problem	Communication	0-2
3	3	Test-taker repairs Abby's misunderstanding of the result.	**(D1)** Monitoring and repairing the shared understanding	Communication	0-2
3	4	Test-taker offers a plan of how to reach the optimum solution.	**(C1)** Communicating with team members about the actions to be/ being performed	Communication	0-2
3	5	Test-taker asks Abby for her point of view before implementing the plan.	**(B1)** Building a shared representation and negotiating the meaning of the problem (common ground)	Communication	0-2
4	1	Test-taker implements the plan as discussed with Abby.	**(C2)** Enacting plans	Action	0-2
4	2	Test-taker shares his or her understanding of the result (fish conditions).	**(D2)** Monitoring results of actions and evaluating success in solving the problem	Communication	0-2
4	3	Test-taker asks Abby for her point of view before implementing the plan.	**(C1)** Communicating with team members about the actions to be/ being performed	Communication	0-2
5	1	Test-taker implements the plan as discussed with Abby.	**(C2)** Enacting plans	Action	0-2
5	2	Test-taker shares his or her understanding of the result (fish conditions).	**(D2)** Monitoring results of actions and evaluating success in solving the problem	Communication	0-2
5	3	Test-taker asks Abby for her point of view before implementing the plan.	**(C1)** Communicating with team members about the actions to be/ being performed	Communication	0-2
6	1	Test-taker implements the plan as discussed with Abby.	**(C2)** Enacting plans	Action	0-2
6	2	Test-taker shares his or her understanding of the result (fish conditions).	**(D2)** Monitoring results of actions and evaluating success in solving the problem	Communication	0-2
6	3	Test-taker asks Abby for her point of view before implementing the plan.	**(C1)** Communicating with team members about the actions to be/ being performed	Communication	0-2
7	1	Test-taker provides reflective feedback on the work with Abby.	**(D3)** Monitoring, providing feedback and adapting the team organisation and roles	Probe, MC	0-2

Note: Score points are assigned based on exhibiting behaviour (performing actions or communicating). Items are scored polytomously (0, 1, 2) according to levels of competency.

Sample CPS unit: CLASS LOGO

Unit classifications

Context: in-school | outside school
Contents: consensus building, win-win negotiation, hidden profile (jigsaw) task
Type of CPS task: decision making | co-ordination | production
Number of agents: three agents, including the student
Target unit timing: 5 minutes | 10 minutes | 15 minutes | 20 minutes

Unit overview (team composition, problem context and overview of tasks)

In this unit, a team of three students – the test-taker, Mark and Sarah (two computer agents) collaborate to produce a logo for a sports event. The goal is to achieve a five-star rating from the class. Mark and Sarah draw the logo and the test-taker's role is to lead the group.

The first task of the unit is an initial discussion between the test-taker, Mark and Sarah on how to design the logo. The team then produces drafts that are rated. The test-taker encounters challenges in collaborating with Mark and Sarah during this stage. Finally, the test-taker gives feedback on the collaborative tasks. The test-taker is told that the number of attempts to design the draft logo (known as "trials") is limited to five. The underlying structure of the task forces the test-taker to be involved in at least two trials to achieve a five-star rating in order to provide sufficient data for CPS measurement.

Agent overview

Mark represents collaboratively orientated agent behaviour (e.g. he initiates ideas, builds consensus, responds to, supports and praises the test-taker). He also reveals information about what to do in the task (e.g. shares his past experience that is relevant to the task). However, in some situations, Mark shows a misunderstanding of the results. As long as the test-taker repairs any misunderstandings or points out the advantages or disadvantages of different strategies, Mark is persuaded. However, if the test-taker doesn't clarify or repair misinterpretations of results or provide evidence that counters a suggested strategy, Mark will press for a rationale for accepting the strategy. Sarah represents the behaviour of a low collaboratively orientated agent (e.g. she interrupts other members of the team, disagrees with the test-taker and Mark, comments negatively about Mark's work, and doesn't follow plans).

Collaborative problem-solving skills

In this unit, the test-taker demonstrates CPS proficiency by establishing a shared understanding of the problem, repairing misunderstanding, monitoring the agents' work, and building consensus with team members. The specific cells addressed in the framework skills matrix (Table 1) are described below.

Introduction and orientation

The unit starts with a briefing on the scenario outline and training on the Chat, Control Panel and Task Space areas of the screen. This section is not timed or scored.

Figure 7.8 [1/2] ▪ **Sample unit CLASS LOGO: Introduction**

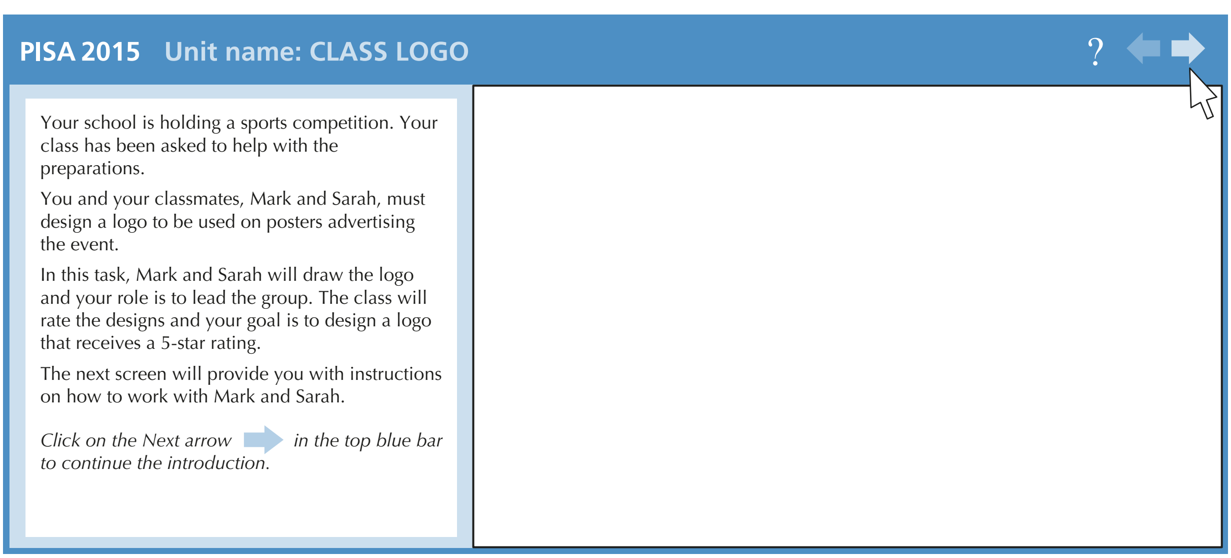
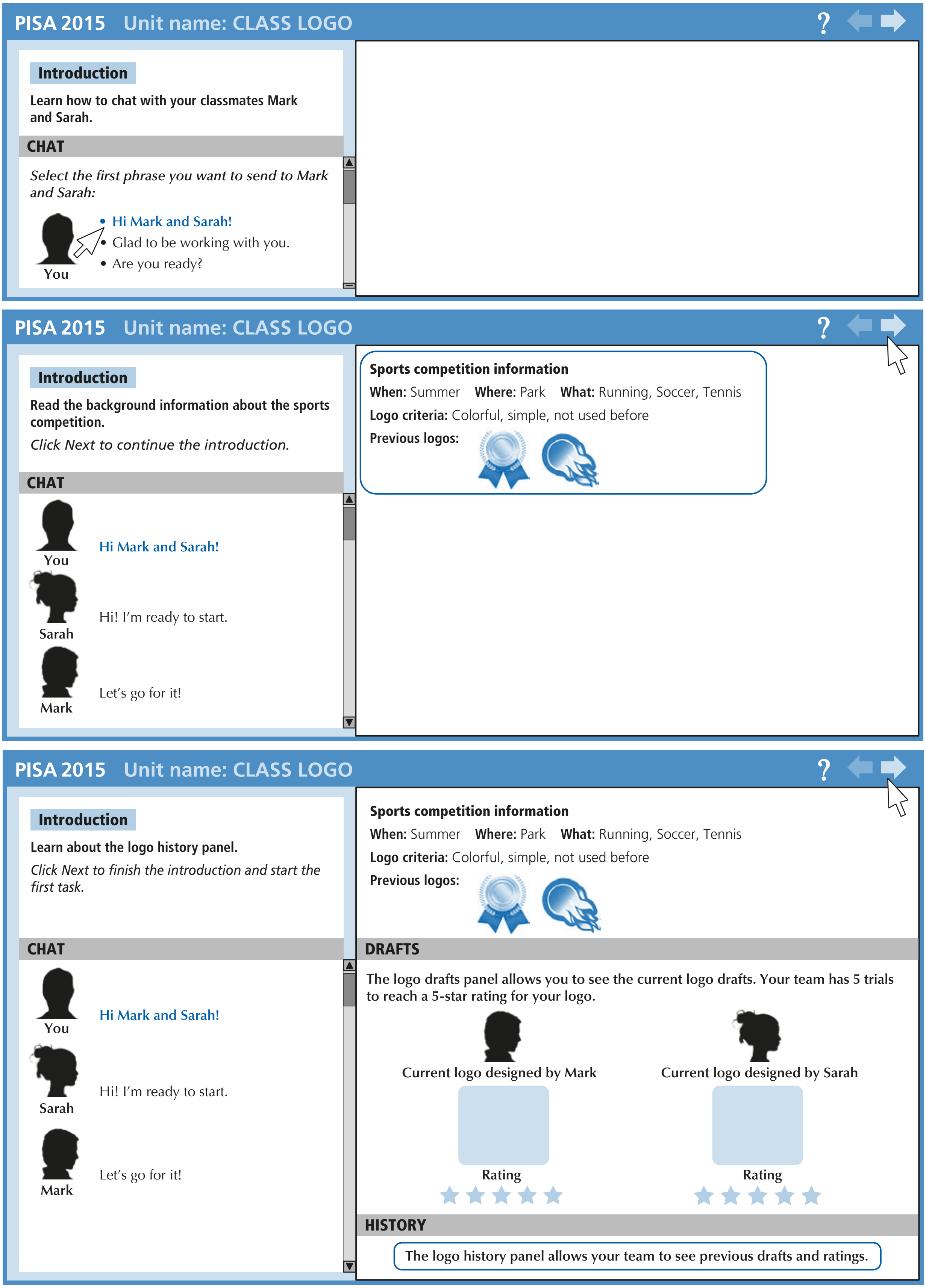


Figure 7.8 [2/2] ■ **Sample unit CLASS LOGO: Introduction**



Outline of unit tasks

Task 1: Establish shared understanding

Activity

- **Item 1:** The test-taker asks Mark and Sarah to describe their abilities in logo design. Mark and Sarah provide a short description. If the test-taker doesn't ask after a certain amount of time or a set number of exchanges, then Mark initiates by describing his ability. There can be multiple exchanges to release the information gradually.
- **Item 2:** The test-taker asks Mark and Sarah about the tools available for them to design the logo. If the test-taker does not do this, then Mark initiates and provides a description.
- **Item 3:** The test-taker offers a plan of how to design a logo (e.g. provides suggestions in the chat on symbols and colours) and asks Mark and Sarah for their point of view. Mark asks the test-taker to provide reasoning (e.g. why do you think so?). If the test-taker provides some reasoning for the plan, then Mark agrees. Otherwise, Mark disagrees and shares his alternative plan with the team. Sarah disagrees with both the test-taker and Mark's plans and suggests her own plan without providing any reasoning.
- If the test-taker doesn't offer an idea, then Mark and Sarah prompt. If still no idea is offered, then Mark and Sarah suggest two different ideas for use of symbols and colours.
- **Item 4:** The test-taker has to ensure that Mark and Sarah agree (e.g. monitor shared understanding) before clicking on "Next" to allow them to produce draft logos. If the test-taker doesn't offer to click "Next", then Mark will rescue and ask if they should do that. When the test-taker clicks "Next", a pop-up asks if all the team members are ready to design the first logo draft. If the test-taker did not agree with Mark and Sarah beforehand, then they can interject here and the test-taker can repair before clicking "Yes" to proceed.

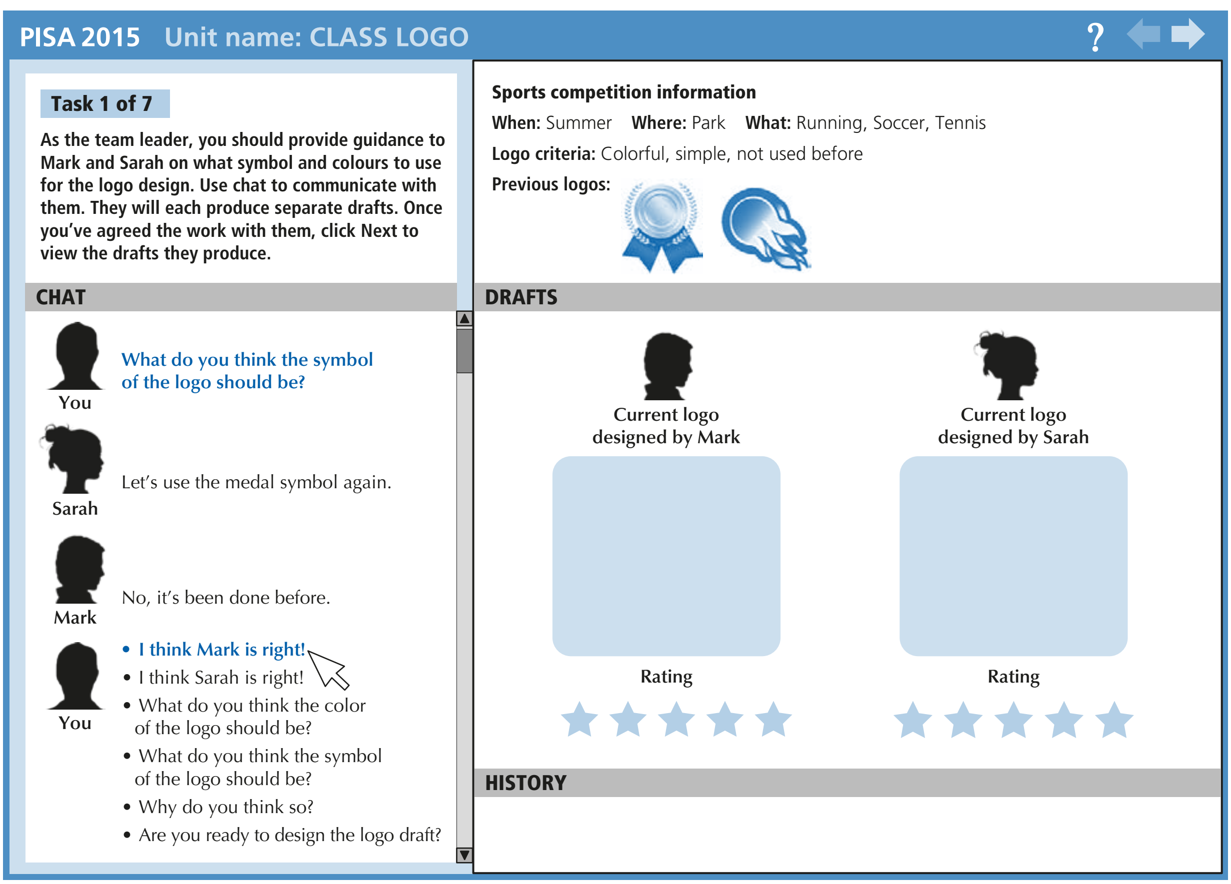
Convergence

A plan is agreed. The test-taker sees Mark and Sarah's draft logos.

CPS skill(s) assessed:

(A1) Discovering perspectives and abilities of team members; **(A2)** discovering the type of collaborative interaction to solve the problem, along with goals; **(C1)** communicating with team members about the actions to be/being performed; **(B1)** building a shared representation and negotiating the meaning of the problem (common ground).

Figure 7.9 ■ **Sample unit CLASS LOGO: Task 1**



Task 2: Monitoring the results and repairing misunderstanding

Activity

- **Item 1:** The test-taker monitors whether Mark and Sarah followed the plan as discussed, and raises additional comments and suggestions to improve the logo drafts.
- **Item 2:** The test-taker asks the agents for their points of view and their readiness to proceed before clicking on "Rate the logos". While Mark is ready to rate the logos, Sarah raises concerns regarding the readiness of the logo drafts, without providing any reasoning. The test-taker asks Sarah to explain her concerns. If the test-taker does not ask, Mark initiates the question. The team agrees to rate the logo drafts.
- **Item 3:** The test-taker shares his or her understanding of the result (the rating and comments for each logo draft). If not, Mark provides a reasonable interpretation.
- **Item 4:** The test-taker has to offer a plan of how to proceed (e.g. "let's change the symbol"). If the test-taker doesn't offer an idea, then Mark can prompt. If still no idea is offered, then Mark will suggest an idea himself.
- **Item 5:** Sarah raises a negative comment regarding Mark's logo draft (e.g. "I don't think that we should work with Mark's logo. It got a very low rating. Let's switch to mine."), but Mark's logo receives a higher rating than Sarah's logo. The test-taker has to repair Sarah's misunderstanding of the collaborative work and/or the results, as well as clarify the roles of the team members.

Convergence

The test-taker can see the ratings and comments for the logo drafts. A plan is decided.

CPS skill(s) assessed:

(D2) Monitoring results of actions and evaluating success in solving the problem; **(D1)** monitoring and repairing the shared understanding; **(B3)** describing roles and team organisation (communication protocol/rules of engagement).

Figure 7.10 [1/2] ■ **Sample unit CLASS LOGO: Task 2**

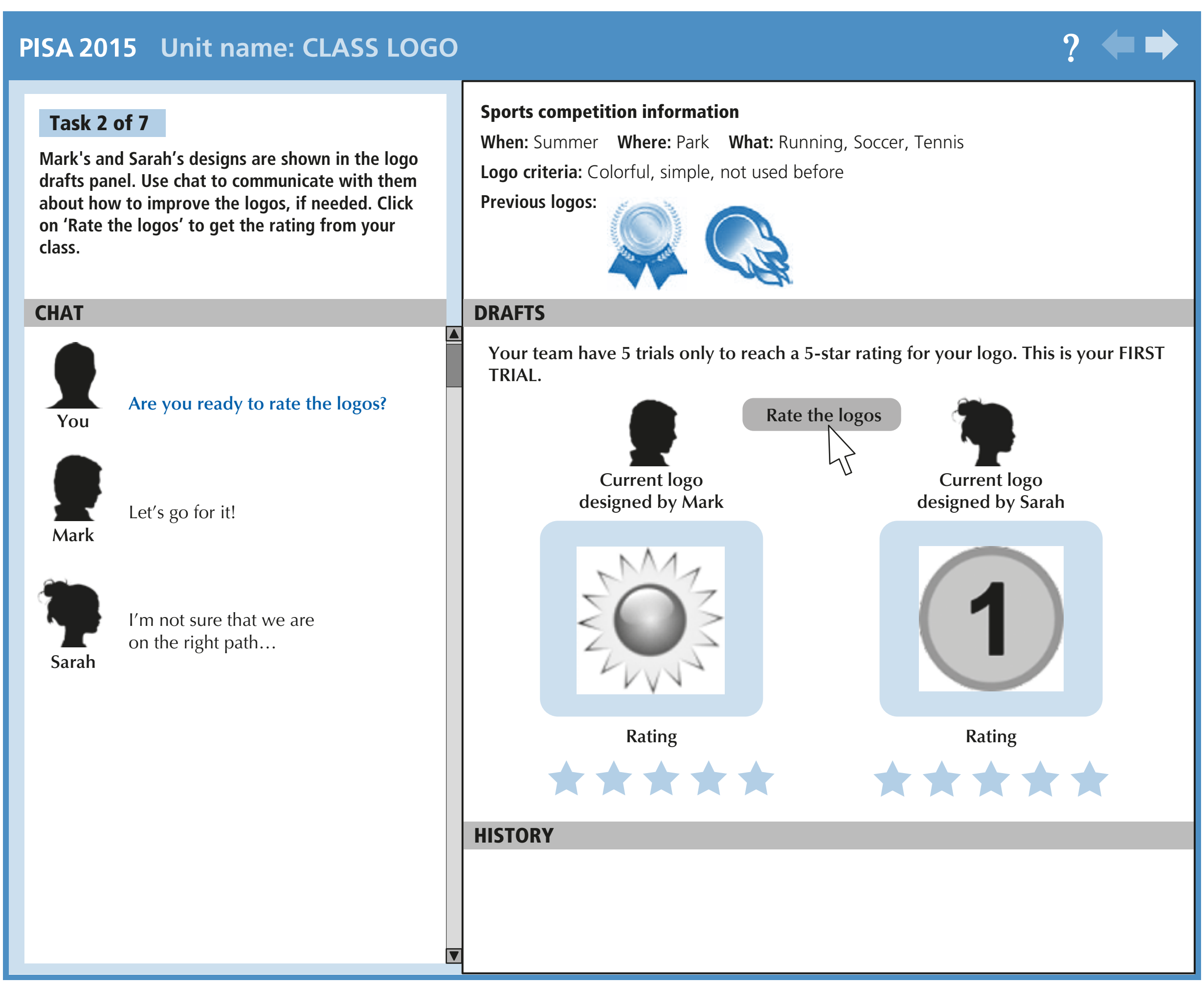
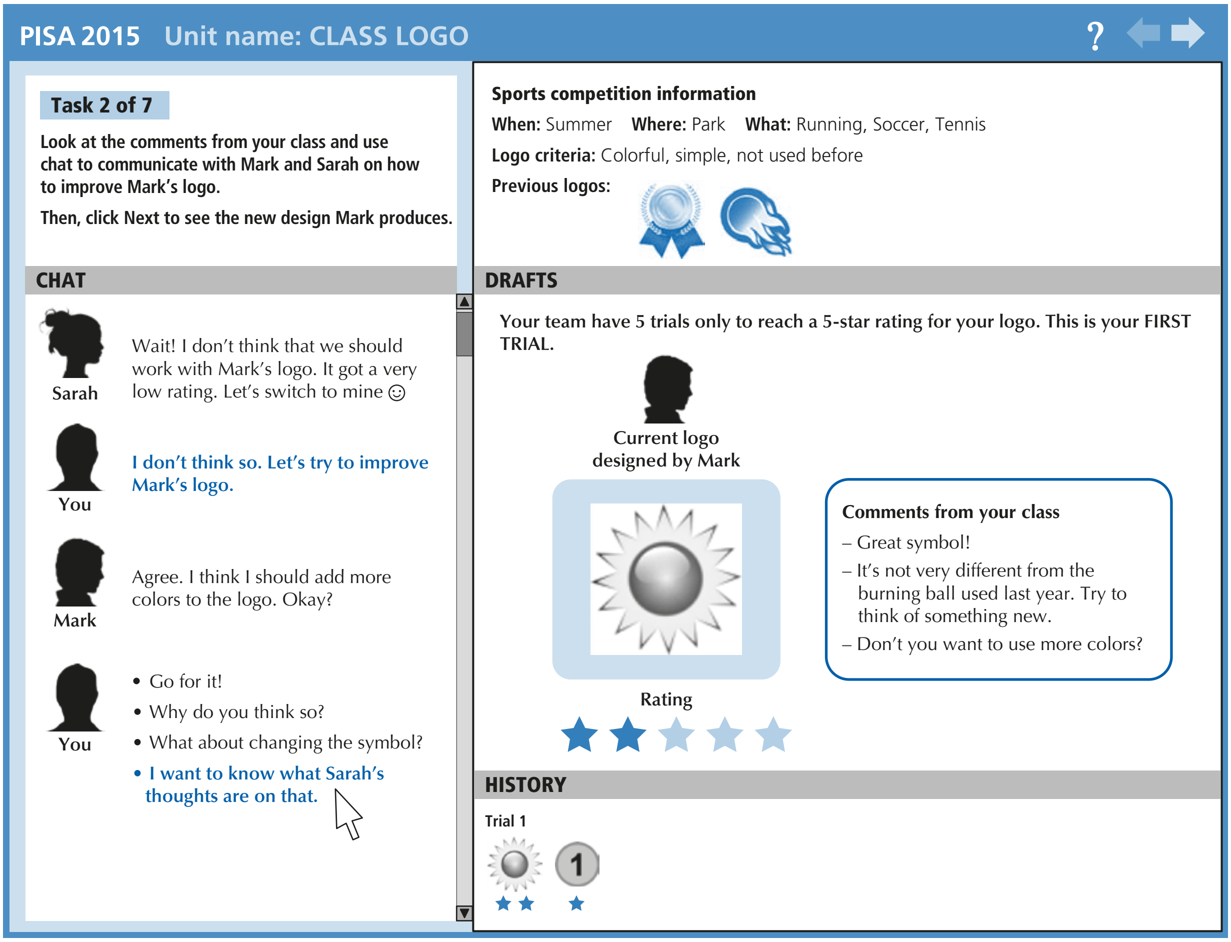


Figure 7.10 [2/2] ■ **Sample unit CLASS LOGO: Task 2**



Task 3: Monitoring and repairing the shared understanding

Activity

- **Item 1:** The test-taker monitors whether Mark and Sarah followed the plan as discussed, and raises additional comments and suggestions to improve the logo drafts.
- **Item 2:** The test-taker discovers that Sarah didn't provide an updated version for the logo as discussed. The test-taker asks Sarah to share the updated draft (e.g. "Sarah, can you share your new draft with us?"). If the test-taker does not ask, then Mark prompts Sarah with a question. Sarah then shares the draft with the team.
- **Item 3:** The test-taker asks the agents for their points of view and their readiness to proceed before clicking on "Rate the logos". If the test-taker does not ask, then Mark initiates the question. The team agrees to rate the updated logo drafts.
- **Item 4:** The test-taker shares his or her understanding of the result (the rating and comments for each logo draft). Mark provides an incorrect interpretation of the result (e.g. "Oh, now the rating is even worse."). The test-taker has to repair this misunderstanding and/or invites Sarah to comment. Sarah comments with a correct explanation.
- **Item 5:** The test-taker has to offer a plan of how to proceed (e.g. "let's change the symbol"). If the test-taker doesn't offer an idea, then Mark prompts. If still no idea is offered, then Mark will suggest an idea himself.
- The team agrees to proceed.

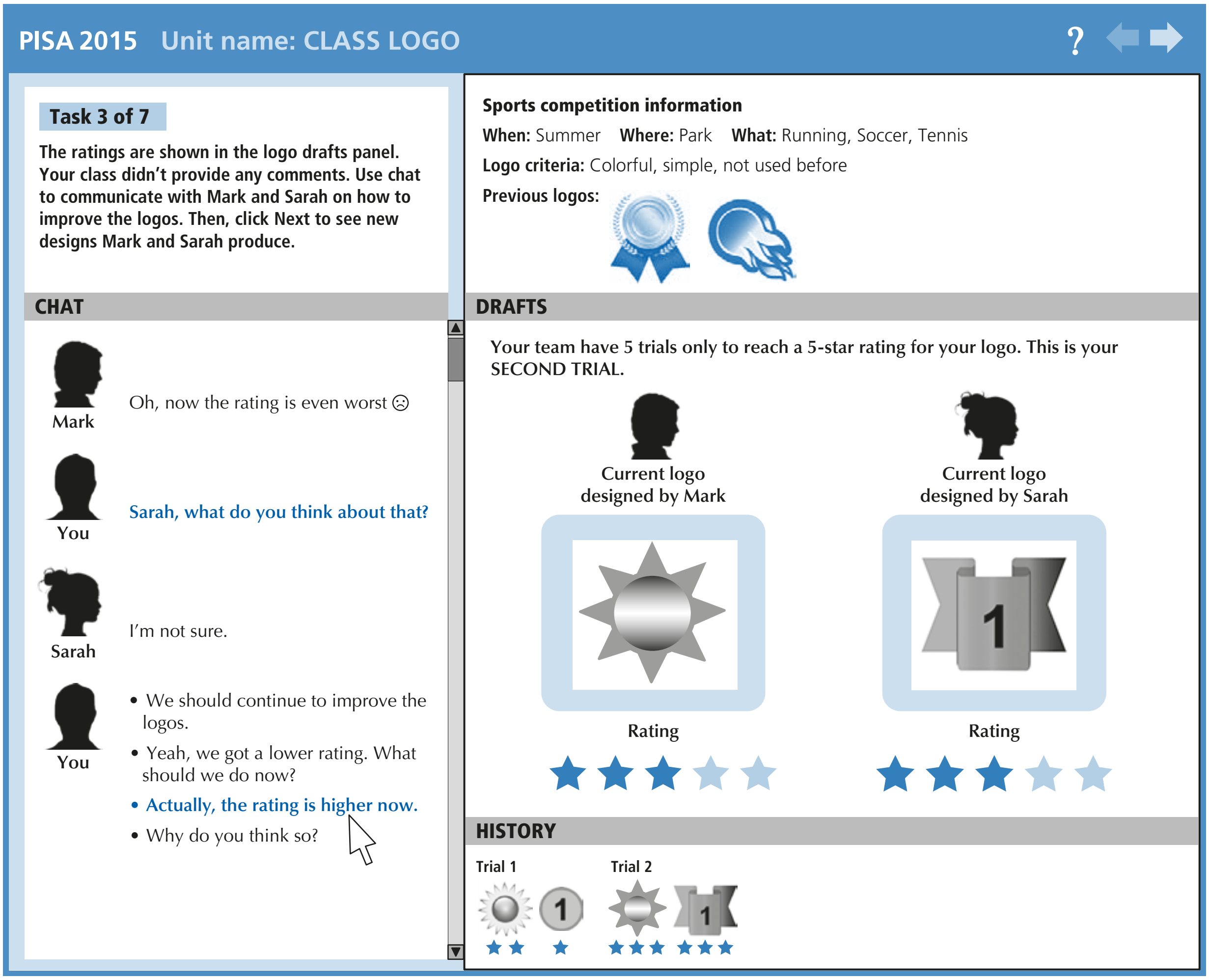
Convergence

The test-taker can see the ratings and comments for the updated logo drafts. Any misunderstanding is repaired. A plan is decided.

CPS skill(s) assessed:

(D2) Monitoring results of actions and evaluating success in solving the problem; **(D1)** monitoring and repairing the shared understanding; **(C1)** communicating with team members about the actions to be/being performed.

Figure 7.11 ■ **Sample unit CLASS LOGO: Task 3**



Task 4: Discovering perspectives and abilities of team members

Activity

- **Item 1:** The test-taker monitors whether Mark and Sarah followed the plan as discussed, and raises additional comments and suggestions to improve the logo drafts.
- **Item 2:** Mark shares with the team that he designed all the previous logos for the class. Sarah comments that it doesn't matter. The test-taker has to explore Mark's newly revealed abilities. Mark provides a clue on how to design a logo that would reach a five-star rating. If the test-taker chooses not to explore Mark's experience, the clue is not presented during this stage.
- **Item 3:** The test-taker asks the agents for their points of view and their readiness to proceed before clicking on "Rate the logos". If the test-taker does not, Mark initiates the question. The team agrees to rate the logo drafts.
- **Item 4:** The test-taker shares his or her understanding of the result (the rating and comments for each logo draft). The test-taker has to offer a plan of how to proceed (e.g. "let's change the symbol"). If the test-taker doesn't offer an idea, then Mark can prompt. If still no idea is offered, then Mark will suggest an idea himself.
- The team agrees to proceed.

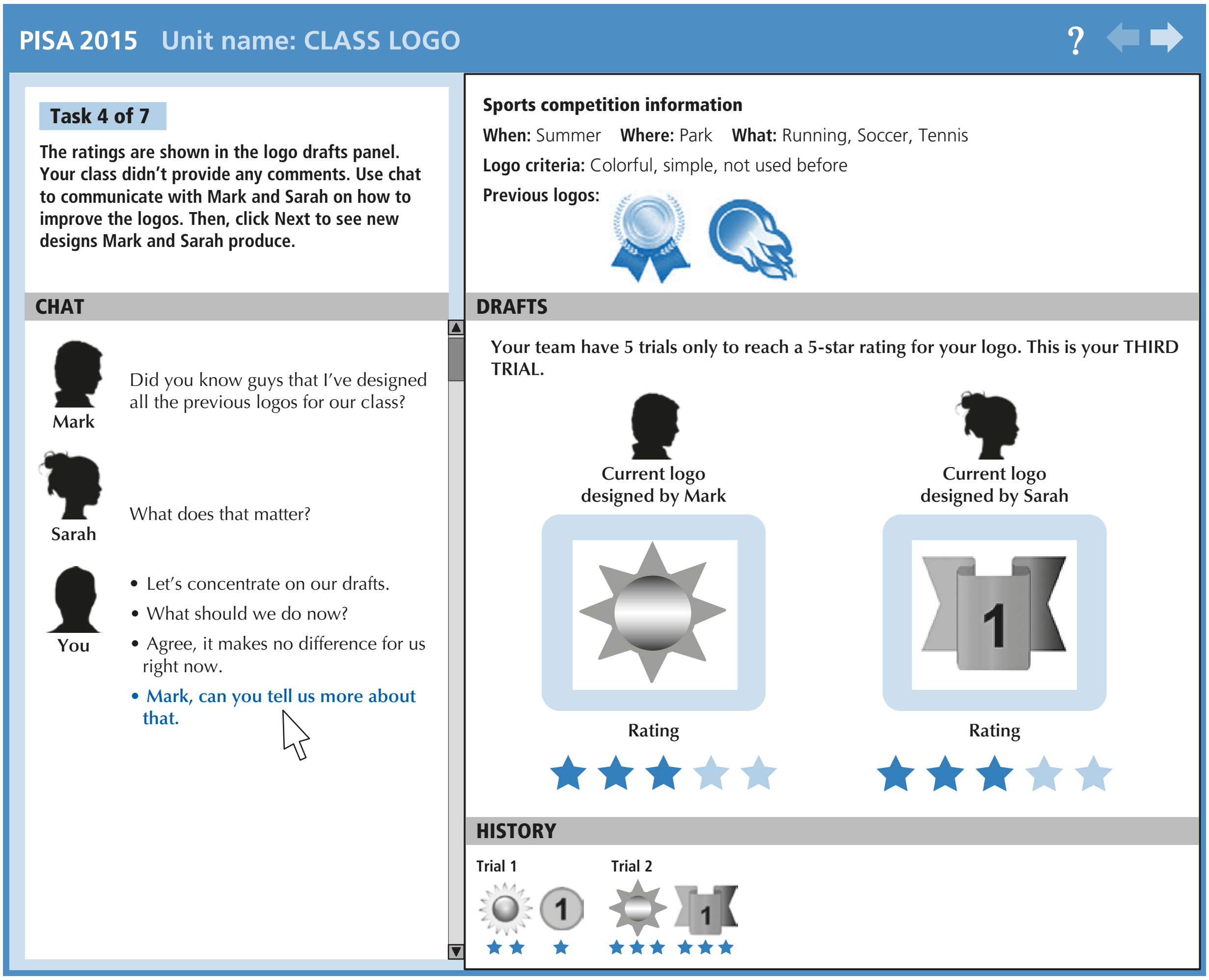
Convergence

The test-taker can see the rating and comments for the updated logo drafts. A clue for a solution is conditionally provided. A new plan is decided.

CPS skill(s) assessed:

(A1) Discovering perspectives and abilities of team members; **(D2)** monitoring results of actions and evaluating success in solving the problem; **(C1)** communicating with team members about the actions to be/being performed.

Figure 7.12 ▪ **Sample unit CLASS LOGO: Task 4**



Task 5-6

These are only presented if applicable, depending on the test-taker's performance.

<u>Activity</u>

Optimising the strategy to solve the problem

- **Item 1:** The test-taker monitors whether Mark and Sarah followed the plan as discussed, and raises additional comments and suggestions to improve the logo drafts.
- **Item 2:** The test-taker asks the agents for their points of view and their readiness to proceed before clicking on "Rate the logos". If the test-taker does not ask, Mark initiates the question. The team agrees to rate the logo drafts.
- **Item 3:** The test-taker shares his or her understanding of the result (the rating and comments for each logo draft). The test-taker has to offer a plan of how to proceed (e.g. "let's change the symbol"). If the test-taker doesn't offer an idea then Mark prompts. If still no idea is offered, then Mark will suggest an idea himself.
- The team agrees to proceed.

<u>Convergence</u>

The test-taker can see the rating and comments for the updated logo drafts. A new plan is decided.

CPS skill(s) assessed:

(D2) Monitoring results of actions and evaluating success in solving the problem; **(C1)** communicating with team members about the actions to be/being performed; **(C2)** enacting plans.

As test-takers may make multiple attempts to optimise the strategy to solve the problem, they would receive scores based on the number of attempts with fewer attempts resulting in higher scores (0-2) for C2. In addition, test-takers would receive the maximum score achieved across attempts for skills D2 and C1.

Task 7: Feedback

Activity

- **Item 1:** The test-taker provides reflective feedback on the work with Mark and Sarah regarding shared understanding of the task.
- **Item 2:** The test-taker suggests a collaborative method (e.g. talk more to Sarah) to promote better collaboration on the task.

Convergence

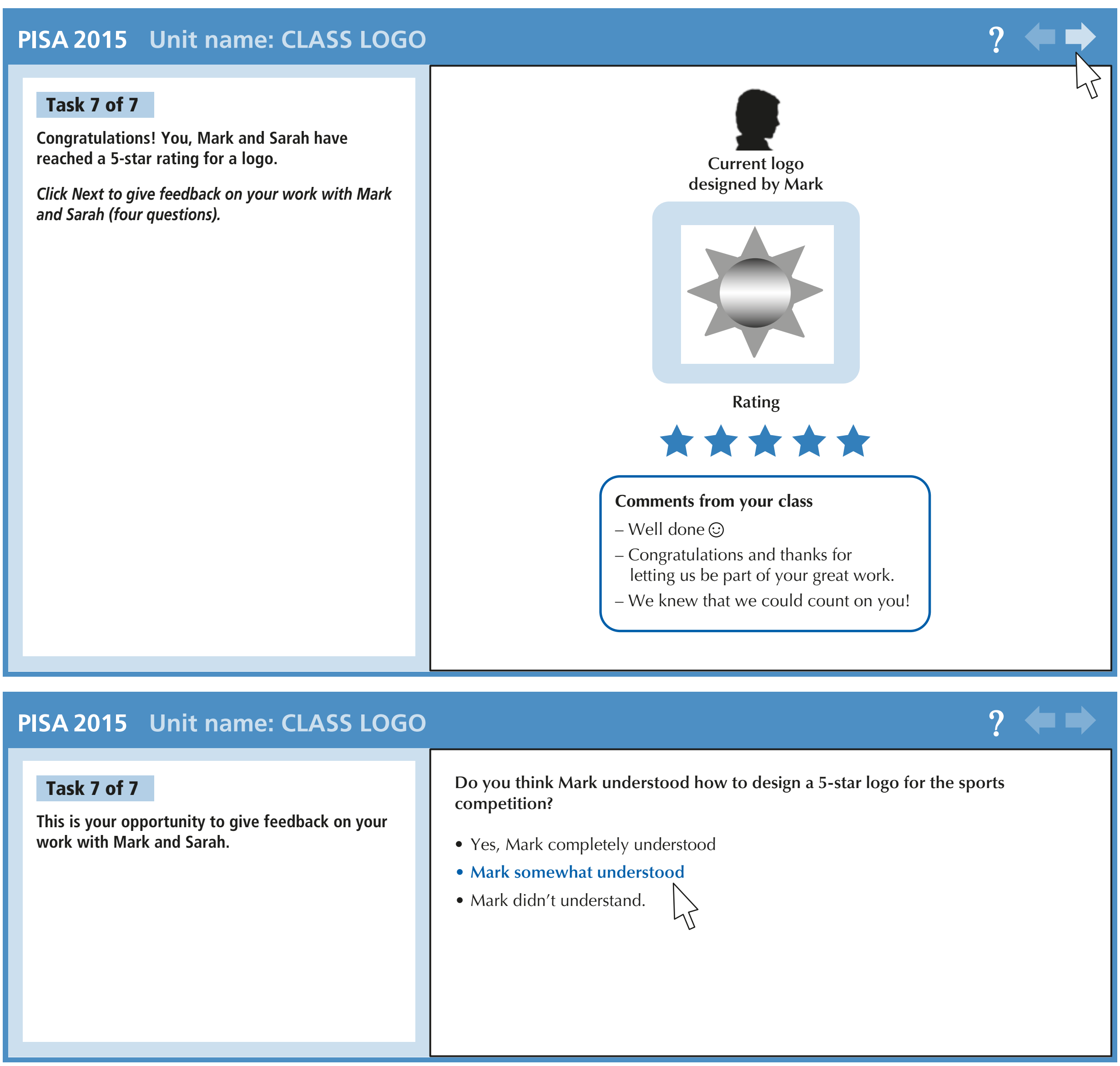
The test-taker, Mark and Sarah share feedback on the task.

CPS skill(s) assessed:

(D3) Monitoring, providing feedback and adapting the team organisation and roles.

These questions are presented in a multiple-choice format. There is a single optimal answer, which receives full credit. Some of the other options would receive partial credit and some options would receive no credit.

Figure 7.13 ■ **Sample unit CLASS LOGO: Task 7**



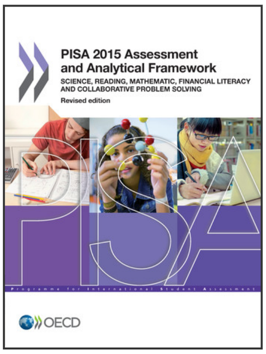
Unit measurement profile

At the end of each task, there is a convergence point. This ensures that all test-takers start from the same point and have the same opportunity to score.

Table 7.12 Profile of assessment items within sample unit CLASS LOGO

Task no.	Item no.	Item short description	Target collaborative problem-solving skill	Data type	Score range (0-x)
1	1	Test-taker explores Mark's and Sarah's abilities in logo design.	(A1) Discovering perspectives and abilities of team members	Communication	0-2
1	2	Test-taker asks Mark and Sarah about the tools available for them to design a logo.	(A2) Discovering the type of collaborative interaction to solve the problem, along with goals	Communication	0-2
1	3	Test-taker offers a plan of how to improve the logo drafts.	(C1) Communicating with team members about the actions to be/ being performed	Communication	0-2
1	4	Test-taker asks Mark and Sarah for their point of view before implementing the plan.	(B1) Building a shared representation and negotiating the meaning of the problem (common ground)	Communication	0-2
2	1	Test-taker monitors whether Mark and Sarah followed the plan as discussed.	(D1) Monitoring and repairing the shared understanding	Communication	0-2
2	2	Test-taker shares his or her understanding of the result.	(D2) Monitoring results of actions and evaluating success in solving the problem	Communication	0-2
2	3	Test-taker repairs Sarah's misunderstanding of the collaborative work and the roles of the team members.	(B3) Describe roles and team organisation (communication protocol/rules of engagement)	Communication	
2	4	Test-taker offers a plan of how to improve the logo design.	(C1) Communicating with team members about the actions to be/ being performed	Communication	0-2
2	5	Test-taker asks Mark and Sarah for their point of view before implementing the plan.	(B1) Building a shared representation and negotiating the meaning of the problem (common ground)	Communication	0-2
3	1	Test-taker shares his/her understanding of the result.	(D2) Monitoring results of actions and evaluating success in solving the problem	Communication	0-2
3	2	Test-taker repairs Sarah's misunderstanding of the actions to be performed.	(D1) Monitoring and repairing the shared understanding	Communication	0-2
3	3	Test-taker repairs Mark's misunderstanding of the result.	(D1) Monitoring and repairing the shared understanding	Communication	0-2
3	4	Test-taker offers a plan of how to improve the logo design.	(C1) Communicating with team members about the actions to be/ being performed	Communication	0-2
3	5	Test-taker asks Mark and Sarah for their point of view before implementing the plan.	(B1) Building a shared representation and negotiating the meaning of the problem (common ground)	Communication	0-2
4	1	Test-taker shares his or her understanding of the result.	(D2) Monitoring results of actions and evaluating success in solving the problem	Communication	0-2
4	2	Test-taker explores Mark's new discovered abilities.	(A1) Discovering perspectives and abilities of team members	Communication	0-2
4	3	Test-taker offers a plan of how to improve the logo design.	(C1) Communicating with team members about the actions to be/ being performed	Communication	0-2
4	4	Test-taker asks Mark and Sarah for their point of view before implementing the plan.	(C1) Communicating with team members about the actions to be/ being performed	Communication	0-2
5	1	Test-taker shares his or her understanding of the result.	(D2) Monitoring results of actions and evaluating success in solving the problem	Communication	0-2
5	2	Test-taker offers a plan of how to improve the logo design.	(C1) Communicating with team members about the actions to be/ being performed	Communication	0-2
5	3	Test-taker asks Mark and Sarah for their point of view before implementing the plan.	(C1) Communicating with team members about the actions to be/ being performed	Communication	0-2
6	1	Test-taker shares his or her understanding of the result.	(D2) Monitoring results of actions and evaluating success in solving the problem	Communication	0-2
6	2	Test-taker offers a plan of how to improve the logo design.	(C1) Communicating with team members about the actions to be/ being performed	Communication	0-2
6	3	Test-taker asks Mark and Sarah for their point of view before implementing the plan.	(C1) Communicating with team members about the actions to be/ being performed	Communication	0-2
7	1	Test-taker provides reflective feedback on the work with Mark and Sarah.	(D3) Monitoring, providing feedback and adapting the team organisation and roles	Probe response	0-2
7	2	Test-taker suggests a collaborative method to improve CPS performance.	(D3) Monitoring, providing feedback and adapting the team organisation and roles	Probe response	0-2

Note: Score points are assigned based on exhibiting behaviour (performing actions or communicating). Items are scored polytomously (0, 1, 2) according to levels of competency.



From:

PISA 2015 Assessment and Analytical Framework

Science, Reading, Mathematic, Financial Literacy and Collaborative Problem Solving

Access the complete publication at:
https://doi.org/10.1787/9789264281820-en

Please cite this chapter as:

OECD (2017), "PISA 2015 collaborative problem-solving framework", in *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic, Financial Literacy and Collaborative Problem Solving*, OECD Publishing, Paris.

DOI: https://doi.org/10.1787/9789264281820-8-en