Hayashi, Y., et al. (Eds.) (2017). Workshop Proceedings of the 25th International Conference on Computers in Education. New Zealand: Asia-Pacific Society for Computers in Education

Validation of Collaborative Problem Solving Process Framework from Evidence of Student Observations for Developing Generic Measures

Nafisa AWWAL[a*], Patrick GRIFFIN[a], Zhonghua ZHANG[a], Claire SCOULAR[a], Monjurul ALOM[a], Daniel JIMENEZ[a] & Mark WILSON[a]

[a]Assessment Research Centre, Melbourne Graduate School of Education, University of Melbourne, Australia

*n.awwal@unimelb.edu.au

Abstract: This study recognises the role of collaborative problem solving (CPS) as an important 21st century skill for a higher quality workforce. This study aims to develop performance measures of individuals CPS ability while being engaged in collaborative problem solving tasks online with another human (H2H). With this aim in mind, the authors propose a new CPS Process Framework. The proposed framework portrays CPS as a consolidation of collaboration and problem solving frames. Indicators of observable behaviours are designed for coding and later scoring those behavioural indicators mapped to this framework. Empirical data including student observation will be used as evidence to explore the validity of this new framework. It is expected that the use of multiple data and phases of analysis will enable and provide insight how CPS processes among H2H dyads evolve during CPS assessments in an online collaborative environment.

Keywords: collaborative problem solving, CPS, collaboration, computer-supported, collaborative, problem solving, process, peer interaction, CPS framework, student observation, log file, process data

1. Background of the Study

Collaborative problem solving (CPS) is now well recognised in industry as a core competency of today's knowledge economy and has taken a central role in recent theoretical and technological developments in education research. It is a relatively new research area and its concepts, methods, and research ideas link collaborative learning, problem solving, data mining, and psychometrics (Kozma, 2009). The OECD decision to assess CPS in the Programme for International Student Assessment (PISA, www.oecd.org/pisa) in 2015 and the pioneering work in Assessment and Teaching of 21st Century Skills study (ATC21S™, www.atc21s.org) has stimulated interest in CPS research as a 21st century skill suitable for formative assessment. The construct has been situated in the zones of education, psychology and employment, often in the context of discussion of 21st century skills. According to ATC21S study, 21st century skills did not all need to be new (Griffin, 2012; Griffin, McGaw & Care, 2012), rather it was argued to be those that must be brought to bear in today's worlds of education, living and work for individuals to function effectively as students, workers and citizens. Collaborative problem solving (CPS) combines critical thinking, problem solving, communication and collaboration (Griffin & Care, 2015). CPS is a joint activity where groups execute several steps to transform a current state into a desired goal state, in which a group may require varied knowledge, expertise and skills, both in terms of interpersonal dynamics as well as in cognitive processes, which is unlikely to be possessed by any one individual. Where a concept has this level of complexity, tasks designed to measure the construct may present challenges both in terms of how a group of individuals might approach that task as well as in terms of what processes the individuals might use to contribute to the resolution of the task, and in terms of finding ways to observe the characteristics in an unconfounded way. CPS thus can be truly be recognised as the "new smarts" in both the assessment and learning domain.

From the outcome of these two pioneering studies, it became evident among researchers that a change to coding and scoring of such complex assessment is required. During the ATC21S experience, repeated attempts to scale score performances using the Hesse et al. (2015) framework have been frustrating and expensive because each time a new CS task was developed the whole process of coding, calibration and scoring was required to be done separately. There is a desperate need for generic indicators that would cover multiple tasks and more than that, enable new tasks to be developed to automatically generate the new indicators. This has not been true for either PISA or ATC21S CPS tasks. The research (Griffin, Care & Wilson, 2015) under the Australian Research Council Grant is focused on identifying efficiency of scoring as a response to the two crucial deficiencies identified in the previous model. The first is a lack of generic scoring encoding procedures and the second is a matter of group size and the differences are of ability between individuals within a group. This paper addresses the first of those issues – issue of a generic scoring encoding system. The answer to this issue was not perceived in the PISA structure, since PISA tasks examined individual persons within a group collaborating to resolve the problem space. As a consequence, task design is implicated in that with four people in a group resolving the problem, tasks need to be designed with unique contribution for four people. This primary focus of this paper is to identify generic coding system which remained overlooked in both ATC21S and PISA.

2. CPS Process Framework

The theoretical understanding of this new framework (Griffin et al., 2015) is derived from observation of people resolving collaborative problem solving. Both PISA and ATC21S lead on from previous theoretical concepts (O'Neil, 1999; OECD, 2012; Polya, 1957) in defining collaborative problem solving within an educational setting. But the result was not efficient. In PISA conceptual framework the dimensions of collaborative problem solving move away from Polya's mathematical problem solving model into a more exploratory and undeclared complexity of problems. Although sharing some similarities with the PISA model, the new framework developed here is different. It's based on direct observation of people solving problems and explanation of the process why and how they were solving the problems and the way they did, including how they collaborated with their partners. Collaboration itself became more clearly defined as a result of this process to indicate it is a combination of a single shared goal, participants being able to make a unique contribution to the problem resolution, a capacity of people to depend upon each other, and a realisation that each member benefited from the work and contribution of other members. This clearly separated from teamwork and from such things as the PISA human to agent model (H2A) and even from the ATC21S human to human model (H2H). The decision to test the direct observation using the steps of exploring defining, planning, implementing, evaluating and reflecting on the process and structure of common goal, dependence, benefit, and contribution meant that a new matrix similar to that of the PISA was derived. There were essential differences which are explored in this paper.

Table 1 presents the new theoretical CPS Process Framework proposed. Each of the boxes in the framework represents the criteria for identifying the demonstration of the indicator within each capability. It will be clear from description in the following sections that PISA's dependence upon Polya and its necessity for linking it to their 2012 individual problem solving meant that it was compromised in terms of its scoring and capabilities. Alternatively, ATC21S framework was based upon collaborative learning and computer-assisted collaborative learning but not on collaborative problem solving. It was a worthwhile addition to the initial configuration of collaborative problem solving that the researchers of this study believe the need for a new model.

2.1. Capabilities

The proposed framework describes the CPS process as consisting of six capabilities Exploring, Defining, Planning, Implementing, Evaluating, and Reflecting.

Exploring refers to participants searching and probing both the social and problem space in a task for building an understanding and perception of the problem. During this process, individuals must deal with queries like "*What do we have?*". It is assumed that individuals' initial reaction to a problem is

likely be to engage and explore the task space to familiarise themselves and to build an understanding of the problem. Both their actions and role within a task could guide their understanding on the importance how their own contributions could achieve success in the given task. In addition, the ability to interact with their partners and realising the need for such interaction is expected to support in their success. There is less coverage seen for this capability from other available frameworks.

Defining focuses on students jointly outlining the problem. In this process, individuals will find answers for "*What is the problem?*". In real-life scenarios, problems are often vague. For good collaboration in like conditions, it is vital to establish a shared vision of the problem (Barron, 2000). To achieve such a common ground, individuals need to identify any gaps in their understanding through managing own resources; sharing, requesting and interpreting information received, and integrating resources to build their mutual understanding of the problem and what is required to solve it (Dillenbourg, 1999; Hesse, Care, Buder, Sassenberg, & Griffin, 2015). This capability extends familiarising beyond self by including others to build and maintain a joint understanding. The idea of this capability is well informed by existing literature of both PS and CPS.

Planning is the process of deliberating a prearranged course of actions or set of steps required to accomplish a certain goal or target (Hayes-Roth & Hayes-Roth, 1979) while at the same time revealing students' ability to develop strategies based on the steps required to solve the problem (Miller, Galanter, & Pribram, 1986). At this stage, individuals will likely ascertain "*What is the plan?*". For planning, individuals need to address a shared problem representation by organising information, analysing the problem and setting a goal to provide the basis for a coordinated solution and to formulate hypotheses for stages of steps required in achieving the desired joint goal (Hesse et al., 2015; Weldon & Weingart, 1993). Researchers consider this capability crucial in solving problems whether independently or collaboratively (Hayes-Roth & Hayes-Roth, 1979), and has been reflected throughout many research studies.

Implementing refers to the way in which individuals approach collaborative tasks and their execution of plans for solving the problem. In this process, individuals join force to utilise their knowledge and expertise to test their hypotheses and execute plans from their previous planning process. Here the focus of individuals is to find "*How do we implement our plan?*". The focus of this capability is mainly selecting appropriate actions for setting their join plan transferred into action. For a better collaborative work during this phase, individual participation and contributions are perceived as pre-requisite characteristics. This process has similarities with PISA's collaborative component "*taking appropriate action to solve the problem*" which refers to the joint effort of individual to act and follow appropriate steps to solve the problem (OECD, 2013); and with Polya's and PISA's problem solving step "*carry out the plan*" and "*executing*" respectively.

Evaluating focuses on the shared progress of the problem throughout the task. In this context, individuals are required to periodically evaluate their progress throughout their CPS journey to identify what is working and what is not, recognise any deviances from agreed plan, and rectify misunderstandings before they impede their joint work (Dillenbourg and Traum, 2006; Roschelle & Teasley, 1994). During this process, individuals may review "*How did we do?*". Checking progress at different stages of CPS can provide collaborators helpful feedback for forcing necessary adjustments and shaping their future activities. Researchers believe that this process is critical to collaboration (Roschelle & Teasley, 1994), as understanding evolves. In evaluating progress individual are thought to be able to identify connections between information and use this to inform future steps for both current and other tasks. This process overlaps with some of the existing frameworks.

Reflecting refers to individuals need for manifesting both their own and others understand to ensure they are aligned. Here individuals would contemplate on "*What do we learn?*". While reflecting individuals may consider if alternative approaches to a problem are more suitable, whether attempted solutions are appropriate, and revisiting initial hypotheses and assumptions (OECD, 2013). If adaptions or modifications is required, individuals may return to the joint planning stage to reorganise information, alter hypotheses, amend plans or set alternative goals. This process has received almost no coverage in existing PS frameworks, but has similarities with PISA's collaborative component "*monitoring and reflecting*" which refers to the joint effort of individual to act and follow appropriate steps to solve the problem (OECD, 2013).

Table 1. Theoretical CPS Process Framework (Griffin et al., 2015)

Criteria							
Indicators	**E.**	Examines shared resources	Agrees on definition of problem (N)	Develops plan together; Allocate roles (N)	Follows sequential action steps of plan (N)	Evaluates plan (N); Agrees on what they have finished; Returns to planning if solution has not been reached	Implements alternative approaches together
	D. Depend	Asks others questions; Asks others about their resources	Making adjustments to what is relevant and what is not (N)	Agrees to plan (N)	Asks for feedback/contribution from others; Takes turns to identify outcomes of trialling	Negotiates finishing; Develops a common judgement of the outcome	Discusses alternative approaches
	C. Benefit	Take and uses others resources; Responds to others questions	Identify others resources that are useful (N)	Discusses the plan (N)	Integrates other contributions into own actions (N)	Suggests they both finish; Integrates others' evaluations	Asks others for feedback on task outcome
	B. Contribute	Give own resources / information to others; Describes own resources to others	Tells others relevant parts of the problem	Suggests plan	Directs others; Reporting to others	Tells others they have finished; Tells others results of evaluation	Tells others task outcome
	A. Focus (Independent)	Engage with own resources	Identifies parts of the problem	(Trialling resources)	Identifying own outcome of plan	Decides you've finished; Evaluates own choices	Reviews task before completing
		1. Exploring	**2. Defining**	**3. Planning**	**4. Implementing**	**5. Evaluating**	**6. Reflecting**
Capabilities							

2.2. Proficiency Levels

The proposed framework in addition to its six capabilities, is inclusive of five proficiency levels: Focusing, Contributing, Benefiting, Depending, and Metacognitive. These levels are contemplated at varying levels of proficiency across each of the six capabilities. Focus represents the lowest proficiency level whereas metacognitive is considered the highest level of proficiency. In focus level individuals work independently demonstrating very little, if any, collaboration, but are focused on their own tasks. In metacognitive level individuals demonstrate meticulously constructed actions that will likely enhance activities in achieving the goal. Levels of proficiency of individuals may vary based upon the capability that is being measured. For example, an individual may demonstrate as 'Depending' during Exploring, but exhibit less proficiency as 'Contributing' while Reflecting. It is assumed that the most proficient collaborative problem solvers would demonstrate Metacognitive levels across all the capabilities.

3. Methods

3.1. Participants

The research participants (n=20 students) were students of Year 9 from a secondary school in Victoria, Australia. The students were randomly assigned into their dyad pairs (p=10 pairs).

3.2. The Tasks

In this study, student pairs completed one bundle of assessment online developed at the Assessment Research Centre at the University of Melbourne during the ATC21S project (Care, Griffin, Scoular, Awwal, & Zoanetti, 2015; Griffin & Care, 2015) for formative assessment of mapped to the CPS framework (Hesse et al., 2015) and is based on human-to-human (H2H) approaches to assessing CPS. In the tasks, student pairs are given a unique subset of resources and information required to solve the problem jointly. Students must rely on their partner to fully comprehend the problem space and to identify all necessary resources to solve it (Care et al., 2015). The communication between the dyads takes place via free form chat interface.

The bundle used in this study comprised of three tasks, lasting approximately 30 minutes. During the tasks, student pairs (A and B) were seated back to back in the same classrooms to ensure that the only means of communication was the chat interface. In this study, the bundle comprised the following tasks (see Care et al., 2015): "Laughing Clowns", which is content-free task, and "Plant Growth" and "Balance Beam", which are content-dependent tasks.

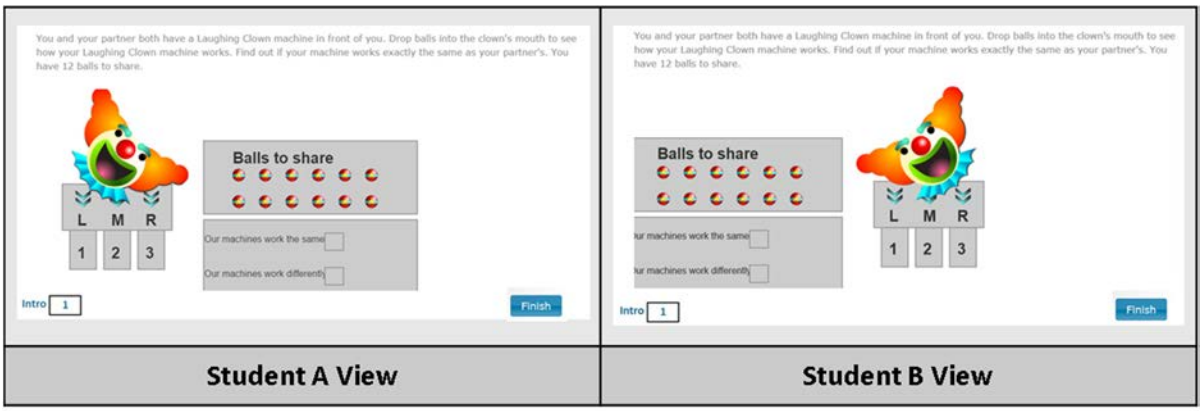

Figure 1. Screenshot of the Laughing Clowns task (showing both individuals' perspective).

The first task, Laughing Clowns, from the administered bundle is the focus of this paper. This task has been designed as symmetric (i.e. both individuals in a collaborative pair are presented with

same information and resources, in other words, same stimulus content and actionable artefacts within the online task space), whereas the other two are asymmetric (i.e. individuals in a pair is presented with different information and actionable artefacts). In Laughing Clowns task, two participants are presented with a clown machine each and 12 balls to be shared with them. The goal for them is to determine whether their clown machines work in the same way. For this to be accomplished, both need to share resources and negotiate how many balls should each use, find patterns, discuss and form rules, and consent on a decision. The students must place the balls into the clown's mouth while it is moving to determine the rule governing the direction the balls will go (Entry: Left, Middle, Right, and Exit= position 1, 2, 3). Each student must then indicate whether they believe the two machines work in the same way (see Figure 1). Students do not have access to each other's screen, so without communication and sharing information are unable to determine the rule governing the other's clown machine.

3.3. Data Collection

3.3.1. Process Data: Log file

In the Laughing Clowns task, there is only a handful of activities is possible for students including the feature to drag any ball, to stop dragging, to drop any ball into their clown's mouth, and to check or uncheck a box to indicate decisions on how their machines worked. Apart from these provisions that are unique to this tasks, a few other common events applicable across all the task in the bundle include indications of the beginning and end of a task, system confirmation messages of individuals' actions, navigational system messages for multiple page tasks, and free-form chat messages for communication with partners. Data for each event is recorded automatically as a single row in a *log file* (records of student–task interactions) and tagged with corresponding student identifier, task identifier, page identifier and role allocation of the acting student in the collaborative session with time-stamping and appropriate indexing (see Table 2). All activities and interactions that are possible within the assessment environment, if recorded systematically as a session log file, can provide salient solution processes in an unobtrusive way (Bennett, Jenkins, Persky, & Weiss, 2003; Zoanetti, 2010). These recorded detailed interactions between the problem solver and the problem environment can be linked to level of proficiency and used to evaluate the process and efficiency with which problem solvers complete games (Pelligrino, Chudowsky, & Glaser, 2001; Williamson, Mislevy, & Bejar, 2006). Individuals' activities in a collaborative session generated log file and patterns in these data were used to assess individuals with the scoring based on their interactions with each other (e.g. occurrence of chat to collaborate etc.) and the task environment (e.g. movement of artefacts etc.). Evidence from the log file indicates activities between the collaborating partners and indicates the level of participation from each to elicit their proficiency level (Awwal, Alom, & Care, 2016). Although not used for this paper, data in the log file also get automatically coded by the scoring engine on Rasch-model as indicators of CPS, producing information on individuals' social and cognitive skill levels (Adams et al., 2015).

Table 2. Excerpt from the log file for the Laughing Clowns task.

736785	auvmir0047	auvmir0047a	6	103	B	0	request-page	1				2016-08-03 16:02:19
736786	auvmir0047	auvmir0047b	6	103	A	0	request-page	1				2016-08-03 16:02:20
736787	auvmir0047	auvmir0047b	6	103	A	1	chat	put a ball in now				2016-08-03 16:02:53
736789	auvmir0047	auvmir0047a	6	103	B	1	move-resource	ball5	L	B-L	B-L	2016-08-03 16:03:12
736791	auvmir0047	auvmir0047a	6	103	B	1	chat	on L it go 1				2016-08-03 16:03:48
736793	auvmir0047	auvmir0047a	6	103	B	1	chat	you try				2016-08-03 16:04:01
736794	auvmir0047	auvmir0047b	6	103	A	1	move-resource	ball11	L	A-L	A-L	2016-08-03 16:04:05
736796	auvmir0047	auvmir0047b	6	103	A	1	chat	me too				2016-08-03 16:04:13
736798	auvmir0047	auvmir0047a	6	103	B	1	chat	i try M				2016-08-03 16:04:32
736800	auvmir0047	auvmir0047a	6	103	B	1	move-resource	ball4	M	B-M	B-M	2016-08-03 16:04:42
736801	auvmir0047	auvmir0047b	6	103	A	1	chat	ok let me know result				2016-08-03 16:04:49
736804	auvmir0047	auvmir0047a	6	103	B	1	chat	M go 3				2016-08-03 16:05:07
736806	auvmir0047	auvmir0047b	6	103	A	1	move-resource	ball10	M	A-M	A-M	2016-08-03 16:05:15
736807	auvmir0047	auvmir0047b	6	103	A	1	chat	me too!				2016-08-03 16:05:22
736810	auvmir0047	auvmir0047a	6	103	B	1	move-resource	ball3	R	B-R	B-R	2016-08-03 16:05:33
736811	auvmir0047	auvmir0047b	6	103	A	1	chat	try R				2016-08-03 16:05:38
736813	auvmir0047	auvmir0047a	6	103	B	1	chat	R go 2				2016-08-03 16:06:02
736815	auvmir0047	auvmir0047b	6	103	A	1	move-resource	ball9	R	A-R	A-R	2016-08-03 16:06:18

736816	auvmir0047	auvmir0047b	6	103	A	1	chat	my (R) went to 1....				2016-08-03 16:06:33
736818	auvmir0047	auvmir0047b	6	103	A	1	chat	Not the same				2016-08-03 16:06:41
736819	auvmir0047	auvmir0047a	6	103	B	1	chat	our machines are different?				2016-08-03 16:06:54
736820	auvmir0047	auvmir0047b	6	103	A	1	chat	yes				2016-08-03 16:06:59
736822	auvmir0047	auvmir0047a	6	103	B	1	select-choice	machines-same	different			2016-08-03 16:07:03

3.3.2. Student Observations: Screen, Audio and Video Recordings

In this study, the collaborative sessions were both audio and video. In addition, students' screen activities were captured for mouse operations and chat discussions along with the recorded tapes during these assessment sessions. The sessions were held at the University of Melbourne in Science of Learning Research classroom that is equipped with such state of the art facilities. The video recordings captured both the students' face as well as all activities on their screen. Students were probed with "Concurrent Oral Reporting", where researchers prompted them strategically for simultaneous commentary, without causing distractions during the completion of the task or inadvertently leading them to any problem solving approach (Ericsson & Simon, 1993). These cues were recorded in the transcripts but not used for any analysis, as students were less verbally responsive during those cues.

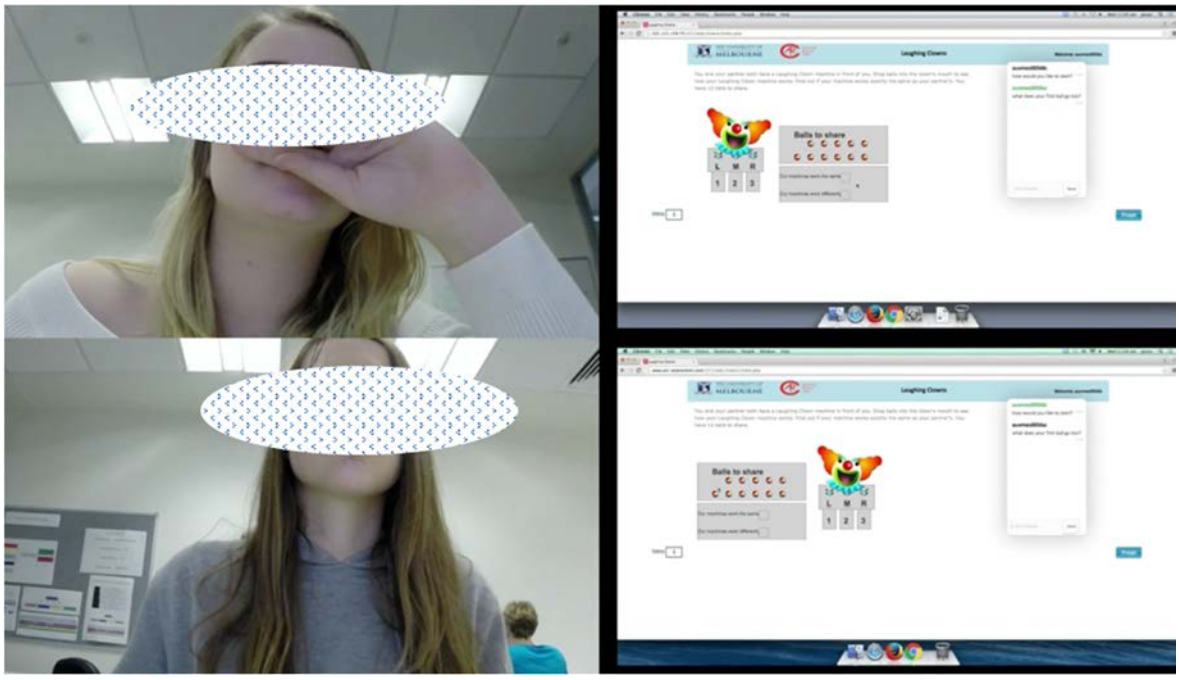

Figure 2. Example of video recording of two students working on the Laughing Clowns task

3.4. Coding and Scoring

A cohort of students was observed while completing the tasks and were scored using the criteria in the theoretical CPS framework (i.e. taking notes on the actions observed for each box in the matrix). This data is analysed using a Guttmann chart. Information identified on the perceived processes undertaken and when will be noted are cross referenced with the log files for verification.

Ten pairs (i.e. 20 students) were video and audio recorded completing one bundle of CPS tasks. An example of the video set up is presented in Figure 2 (students faces have been covered in accordance under our research ethics agreement). Student A can be viewed in the top left quadrant, and their screen perspective in the top right quadrant. Student B can be observed in the bottom left quadrant, with their screen perspective in the bottom right quadrant. In addition to typing their communication to one another in the chat box, they were asked to speak aloud their thought processes as they worked through the task.

The researchers observed the recordings of the collaborative sessions and scored each dichotomously using the theoretical framework. Student chat box communication, actions and speak aloud communication was used to score. A score per criterion was provided for each student across

the whole assessment (all three tasks). For example, where a student was observed describing their own resources to others (Defining/Contribute), they received a 1 in that box, or a 0 if this behaviour was not demonstrated. The researchers discussed the differences in their opinion or observations where appropriate and condensed their scoring into one scoring chart (see Table 3). Each row presents a student, and each column presents a criterion. The numbers in the third row correspond to the coding system presented in Table 1. For example, 1A represents the capability Exploring (1) and the indicator Focus (A). For ease of reference, the criteria descriptions are also presented. Totals for each student and item are provided.

3.5. Data Analysis

The consolidated scoring chart was sorted into a Guttman chart to enable a visual representation of the scoring. The Guttman chart orders student performance according to student demonstrated proficiency, and orders assessment items according to their difficulty (Guttman, 1950). As can be observed in Table 3, scores of red have been highlighted red to assist with visibility. The rows were sorted, according to student total from largest to smallest (top to bottom) so that the most proficient student on the assessment is now at the top, and the least proficient student on the assessment is now at the bottom. In addition, the columns were sorted, according to their totals, from largest to smallest (left to right) so that the easiest item is presented on the left, and the hardest item that is the highest score is on the right.

The modified Guttman analysis allows a qualitative review of the framework and its capacity to be used as a scoring mechanism for CPS assessments. The extent to which the data aligns with the theoretical interpretation of the constructs can be analysed. In addition, the video and audio recording data was triangulated with the log file data. Sections of log files were highlighted from each team that were perceived to be relevant to each of the capabilities in the process. This log file analysis demonstrated evidence of the criteria in the framework. This process has been iterative to inform additional evidence regarding the construct and the framework of CPS.

If the categories and levels illustrated in the Gutmann chart are listed, even with very limited data it is evident that there is a general progression going from the first indicator with steadily rising levels of element. Additional data is required to reinforce the notion that the vertical axis of Table 1 to be forming a hierarchical sequence consistent with the construct of collaboration. On the other hand it does appear to be in hierarchical relationship emerging within each of the five stages of collaboration. More data would be required to test whether this is result of a single dimensional construct. However with more data it could be expected that the construct illustrated in Table 1 appears to be supporting the hypothesis that through developing an independent focus on common goal, the capacity to make an independent contribution; an awareness that there is benefit in what the partners in other collaborators are doing; an acceptance that they depend upon other members of the group and to some extent learn to trust and finally they are able to examine their own thinking in terms of the collaboration. Hence it may well be some beginning evidence of the construct for collaboration but at this stage describing the process of problem-solving has not yet obtained sufficient data to make a conclusion.

4. Discussion and Conclusions

CPS Process Framework is a unique contribution as proposed in the main study by Griffin et al., (2015). The idea presented in the study on the framework is that proficient collaborative problem solvers will begin by exploring both the social and problem space. They are then expected to move forward into sharing their joint resources to develop mutual understanding in defining the problem. Students will then progress in developing a plan together and implement it. Proficient students are then likely follow it up by evaluating and reflecting on the consequences of their results and consider alternative hypotheses where possible. The entire process is possibly repetitive where students may regress to a previous process given the complexity of the imminent activity.

The aim of this study was to present the general idea of the process framework for CPS and present the initial validation done through a series of observations. Using the evidences collected (e.g. log files, recordings, oral reporting and physical observation) researchers could observe student

playing the tasks, identify where they move from one process to the next, then map their judgements to the log files. As an ongoing study, further research is in progress to investigate the validity evidence for this new theoretical CPS Process framework with empirical data.

Table 3. Modified Guttman analysis of the scored categories

	Criteria																													
	Manipulate own resources	Uses only own resources	Develops general plan	Completes task	Give own resources / information to others	Directs others to use resources	Integrates other contributions into own actions	Describes own resources to others	Take and uses others resources	Asks / Enquires about others resources	Agrees on definition of problem	Reports outcomes of actions to others	Describes own resources to others	Shares conflict with others	Examines shared resources	Responds to others questions	Discusses relevant resources before commencing trialling	Negotiates in an attempt to resolve conflict with others	Asks for feedback/contribution from others	Takes turns to identify outcomes of trialling	Tells others task outcome	Discuss joint plan with sequential steps	Negotiates to successfully resolve conflict with others	Asks others for feedback on task outcome	Reviews task before completing	Agrees on task outcomes	Discusses alternative approaches together	Implements alternative approaches	Develops shared action plan with sequential steps	Follows sequential action steps of plan	
	1A	2A	3A	5A	1B	4B	4C	3B	1C	1D	2E	4A	2B	5B	1E	2C	3D	5C	2D	4D	6B	3C	5D	6C	6A	5E	6D	6E	3E	4E	
085b	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	27
086a	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	27
085a	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0	0	26
055b	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	0	1	1	1	1	1	0	0	25
088b	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0	0	0	25
054a	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	24
058b	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	24
090b	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	22
056a	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	0	1	0	0	0	0	0	0	21
090a	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	21
059a	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0	1	1	1	0	0	0	0	0	0	0	20
087b	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	20
088a	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	20
055a	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	0	1	0	0	0	0	0	0	1	0	0	18
054b	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14
058a	1	1	1	1	1	0	0	1	1	0	1	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	11
056b	1	1	1	1	1	1	1	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	10
059b	1	1	1	1	1	1	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9
086b	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	6
087a	1	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6
Total	20	20	20	20	20	17	17	16	16	16	16	15	15	15	15	14	14	13	12	12	11	9	9	6	5	5	4	4	0	0	

Acknowledgments

The work presented here is from the research grant "Measuring individual and group performance in collaborative problem solving" (Griffin et al., 2015) supported under the Australian Research Council's Discovery Projects funding scheme (DP160101678). Acknowledgment for the work presented here is contributed to all the researchers of this project: Patrick Griffin, Mark Wilson, Sandra Milligan, Zhonghua Zhang, Nafisa Awwal, Claire Scoular, BM Monjurul Alom, and Daniel Barrios Jimenez.

References

Adams, R., Vista, A., Scoular, C., Awwal, N., Griffin, P., & Care, E. (2015). Automatic coding procedures for collaborative problem solving. In P. Griffin & E. Care (Eds.), *Assessment and teaching of 21st Century skills: Methods and approach* (pp. 115-132). Dordecht: Springer.

Awwal, N., Alom, M., & Care, E. (2016). Game Design for Learning to Solve Problems in a Collaborative Environment. *Proceedings of the European Conference on Information Management & Evaluation*, 25-33.

Barron, B. (2000). Achieving Coordination in Collaborative Problem-Solving Groups. *The Journal of the Learning Sciences, 9*, 403-436.

Bennett, R. E., Jenkins, F., Persky, H., & Weiss, A. (2003). Assessing complex problem solving performances. *Assessment in Education, 10*(3), 347-359.

Care, E., Griffin, P., Scoular, C., Awwal, N., & Zoanetti, N. (2015). Collaborative problem solving tasks. In P. Griffin & E. Care (Eds.), *Assessment and teaching of 21st century skills: Methods and approach* (pp. 85-104). Dordrecht: Springer.

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

Dillenbourg, P. (1999). *Collaborative learning : cognitive and computational approaches*: Amsterdam ; New York : Pergamon, 1999. 1st ed.

Ericsson, K. A., & Simon, H. A. (1993). Protocol Analysis: Verbal Reports as Data. In. London: The MIT Press.

Griffin, P. (2012). Symposium: The assessment and teaching of 21st century skills. In *The 8th Conference of the International Test Commission Conference*. Amsterdam, The Netherlands.

Griffin, P., Care, E. & Wilson, M. (2015). Measuring individual and group performance in collaborative problem solving. Australian Research Council Grant: DP160101678, Australia.

Griffin, P., & Care, E. (Eds.). (2015). *Assessment and teaching of 21st century skills: Methods and approach*. Dordrecht: Springer.

Griffin, P., McGaw, B., & Care, E. (Ed.) (2012). *Assessment and Teaching of 21st Century Skills*. Dordrecht: Springer.

Guttman, L. (1950). *The Bases for Scalogram Analysis, Chap. 3*.

Hayes-Roth, B., & Hayes-Roth, F. (1979). A cognitive model of planning. *Cognitive science, 3*, 275- 310.

Hesse, F., Care, E., Buder, J., Sassenberg, K., & Griffin, P. (2015). A Framework for Teachable Collaborative Problem Solving Skills. In P. Griffin & E. Care (Eds.), *Assessment and Teaching of 21st Century Skills Volume 2 - Methods & Approach*. Dordrecht: Springer.

Kozma, R. (2009). Assessing and teaching 21st century skills: A call to action. In F. Schueremann & J. Bjornsson (Eds.), *The transition to computer-based assessment: New approaches to skills assessment and implications for large scale assessment* (pp. 13-23). Brussels: European Communities.

Miller, G. A., Galanter, E., & Pribram, K. H. (1986). *Plans and the structure of behavior*. Adams Bannister Cox.

O'Neil, H. F. (1999). Perspectives on computer-based performance assessment of problem solving. *Computers in Human Behavior*(3-4), 255.

OECD. (2012). PISA 2012 Field Trial Problem Solving Framework. In.

OECD. (2013). *PISA 2015: Draft Collaborative Problem Solving Framework.* Retrieved from http://www.oecd.org/callsfortenders/Annex%20ID_PISA%202015%20Collaborative%20Problem%20Solving%20Framework%20.pdf:

Pelligrino, J., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.

Polya, G. (1957). *How to solve it* (2nd ed.). Princeton University Press: Princeton.

Roschelle, J., & Teasley, S. D. (1994). The construction of shared knowledge in collaborative problem solving. *Nato ASI Series of Computer and System Sciences, 128*, 69.

Weldon, E., & Weingart, L. R. (1993). Group goals and group performance. *British Journal of Social Psychology, 32*(4), 307-334.
Williamson, D. M., Mislevy, R. J., & Bejar, I. I. (2006). *Automated scoring of complex tasks in computer-based testing*. New Jersey: Lauwrence Erlbaum Associates, Inc.
Zoanetti, N. P. (2010). Interactive computer based assessment tasks: How problem-solving process data can inform instruction. *Australasian Journal of Educational Technology, 26*(5), 585-606.