

P. JERMANN & P. DILLENBOURG

## ELABORATING NEW ARGUMENTS THROUGH A CSCL SCRIPT

### 1. INTRODUCTION

The CSCL community faces two main challenges with respect to learning and argumentation. The scientific challenge is to understand how argumentation produces learning, that is to discover which cognitive mechanisms, triggered by argumentative interactions, generate new knowledge and in which conditions. The engineering challenge is to determine how to trigger productive argumentation among students. These two challenges are often investigated in parallel, but this contribution focuses on the latter.

There are two ways to favour the emergence of argumentation, either proactively, by structuring collaboration, or retroactively, by regulating interactions (e.g. a tutor monitors the pair dialogues). Structuring collaboration either means scripting collaborative activities or designing a dedicated communication tool. The features of such argumentation tools constitute a central concern of this book. This contribution addresses both forms of structuring. We describe ArgueGraph, a CSCL script encompassed in a web-based environment, and then compare two different interfaces for this environment.

The notion of script enables us to formalize the educational context in which argumentation is expected to appear. A script is a story or scenario that students and tutors have to play as actors play a movie script. Most scripts are sequential: students go through a linear sequence of phases. Each phase specifies how students should collaborate and solve the problem. A phase is described by five attributes: the task that students have to perform, the composition of the group, the way that the task is distributed within and among groups, the mode of interaction and the timing of the phase (Dillenbourg, 2002). The ArgueGraph script fosters argumentation by forming pairs of students with conflicting opinions. Conflicting situations are of particular interest with respect to collaborative interaction because they enable socio-cognitive conflict (Doise & Mugny, 1981): a social conflict (having a different perspective) has to be solved through a cognitive coordination of the points of view. However, further studies showed that conflict is neither a necessary, nor a sufficient condition for cognitive change. Beyond the intensity of conflict, it is the verbalization necessary to solve the conflict, which seems related to learning effects (Blaye, 1988; Butterworth, 1982).

P.JERMANN & P.DILLENBOURG

This chapter reports two experiments based on the same script but using different environments. These experiments confirmed what Suthers (this volume), Baker (this volume) and Veerman (this volume) found in different contexts: the ergonomic features of the communication tool influence argumentation. Some features obviously have a strong impact: audio versus text-based communication, synchronous versus asynchronous, etc. What this book shows is that argumentation is also influenced by more subtle features of the medium. Our experiments revealed some of these features.

Similar experiments are described in Schwarz & al. (2001). The authors are interested in the effect of argumentation on collective and individual arguments. They have their students produce arguments about animal experimentation. They show that students produce more complex and elaborated individual arguments as they interact with other students. They also compare two tools used to construct collective arguments. One is inspired by Belvedere (Suthers, this volume) and graphically represents the relations between pros and contras as well as the conclusion of argumentation. The other is a simple table that holds pros and contras and doesn't explicitly show relations between arguments. The collective arguments are more elaborated when using the argumentative map than when using the table. However, this advantage doesn't show up in the final individual arguments.

This chapter is organized as follows. First we describe the learning activity that was used in the two experiments we made. Then, we present the details of each experiment in two separate sections. Then we compare the results from the two experiments and finally we abstract design factors that have an impact on argumentation.

## 2. THE ARGUEGRAPH SCRIPT

The ArgueGraph script is used in the beginning of a master course on the design of educational software. The two reported experiments were run in the normal course settings that combine co-present interactions and distant learning. The environment was part of Tecfa Virtual Campus, used by our students on a daily basis. The learning objectives are to make students understand the relationship between learning theories and design choices in courseware development. For instance, the students have to learn that the notion of immediate feedback is related to the behaviourist framework and is especially relevant some types of procedural skills acquisition or for rote learning.

## ELABORATING NEW ARGUMENTS

**Question 1**

In a courseware, when a student makes an error, it is better to:

- 1. Tell the student he made a mistake and give him the correct answer.
- 2. Tell the student he made a mistake and give him an indication towards the correct answer.
- 3. Show the student a blinking icon, which allows him to ask the tutor for help.
- 4. Give the student some time to find out the mistake by himself

Argument:

Because the student is stimulated to think about his error| 1

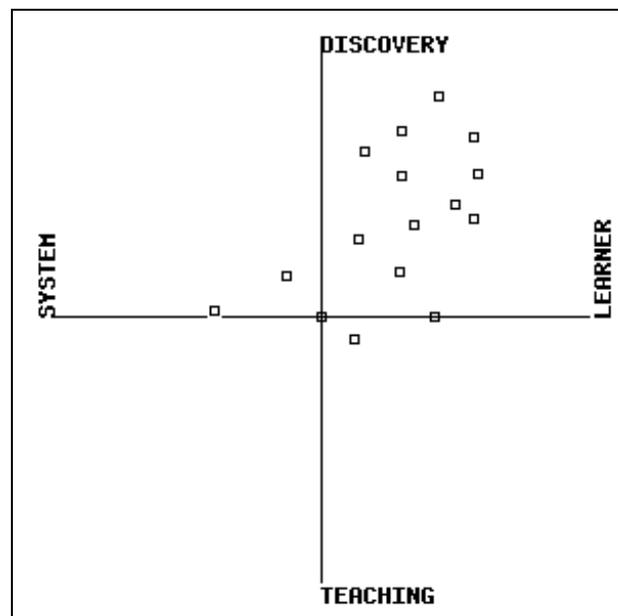
*Figure 1. Question display in version 1. Students use the radio button to make an exclusive choice. They provide a written argument for their choice.*

The script for the activity consists of five phases:

1. **Solo phase:** Students fill in a questionnaire about design principles in courseware development. Questions measure opinions and students provide a short written argument for each of their choices (figure 1). The proposed answers are not correct or incorrect but reflect different pedagogical styles and are grounded in different learning theories to be addressed in this course.
2. **Display phase:** The choices made at phase 1 are then transformed into two scores reflecting whether students privilege system- vs. user-driven interactions on the one hand and a discovery vs. teaching based pedagogy on the other hand. The system draws a scatter plot on the basis of these scores and represents each student's position (figure 2). The choice of scores for each answer was done in a rather arbitrary way. The final sum of scores collected by a student is not a scientific estimate of his pedagogical style. It is only a very rough approximation. The goal is not to produce a very exact value but to use the distances between students to create pairs for the third phase of the script.
3. **Duo phase:** Students fill in the same questionnaire in pairs. Pairs are formed so as to maximize differences within a pair according to the answers the students gave to the individual questionnaire. While filling in the questionnaire the students see the arguments they gave to support their answers in the individual phase. They have to agree on a common choice and provide a common argument.
4. **Debriefing phase:** The system produces a synthesis and a scatter plot representing the "migration" of each student from his initial position to the common position. The synthesis lists the individual and common arguments

given for each question and draws a pie chart with the distribution of answers. Finally, a brief statement presents the relationship between underlying theories and the options the students can select in a question. During this debriefing phase, the teacher reviews all arguments produced by individuals and pairs and relates them to the various theoretical frameworks in the domain (behaviorist, constructivist, socio-cultural, ...). Most pieces of information used in the debriefing have been mentioned in the arguments provided by the students. The teacher's role is to structure this mountain of information into a more coherent framework.

- 5 **Homework phase:** Finally, students get some homework. They have to analyze the answers to one of the questions including thereby theoretical stakes and their own opinion



*Figure 2. Scatter plot of individual answers. Each square corresponds to a student's opinion. The horizontal axis opposes system vs. learner driven interaction. The vertical axis opposes discovery-based learning vs. teaching. Names of students have been removed from the figure for confidentiality.*

This script, as many others (Dillenbourg, 2002) is not purely collaborative, as it includes also individual phases (1 and 5) and collective phases (2 & 4). Only phase 3 is really collaborative. "Collective" differs from "collaborative" due to the fact that it does not necessarily imply rich interactions among students. Simply, the system collects individual productions or data and makes them available to the whole group (Jermann & Dillenbourg, 1999).

## ELABORATING NEW ARGUMENTS

This activity was run twice with different timing of the phases and tool features. In the first experiment, the whole script was held in a four-hour session. Students had to make an exclusive choice in favour of a particular answer and provide a written argument supporting their choice (figure 1). Making an exclusive choice means that students were able to choose only one of the possible answers to the question asked.

In the second experiment, the script span over three weeks, the first three phases taking place at the student's home. Once they completed phase 1, students were told by email who their partner would be for phase 3. They set up a date and a time with their partner and would both dial up to the network at that time and answer the questionnaire together. Students had to make a non-exclusive choice by rating their level of acceptance for each possible answer through the use of a five-point slider. Labels for the sliders were 'Never', 'Rarely', 'Sometimes', 'Often', and 'Always' (figure 3).

### 3. EXPERIMENT 1

#### *3.1. Research question*

Intuitively, the Argue & Graph activity worked very well, as indicated by the high degree of learner involvement during all phases. However, the tool was not subject to a formal evaluation procedure. Nevertheless, we collected all answers and arguments in order to grasp learning mechanisms. Namely, we investigate the differences between individual answers and pairs' answers.

#### *3.2. Subjects*

We ran two pre-experiments, with 15 students in 1997 and 17 students in 1998 respectively, after which the system was improved with regard to various functionalities. The experiment reported in this section was run on October 22nd 1998 with 18 students. Most students were located in the same quarter of the graph (Figure 1). This phenomenon is probably due to the fact that the questions did too clearly reflect the pedagogical values sponsored at TECFA and did not take into account the technical or financial dimensions of courseware development.

#### *3.3. Variables*

The pairing method used in phase 2 aimed at creating conflictual situations, i.e. make pairs answer a question to which individuals gave different answers. Hence, our first variable is the actual frequency of conflict, i.e. the number of times pair members had given different individual answers in solo. For 52% of the answers there was a conflict between the two solos answers. In other words, the method used to form pairs was rather effective. There is some relation between the distance in the graph and the frequency of conflict (the five pairs with a distance of 1 have a

disagreement rate of 38%, while the pairs with a larger distance have a disagreement rate of 52.5%), but the size of the sample is not sufficient to compute a correlation rate.

Since students answer twice to the same questionnaire, we have the opportunity to examine what differentiates answers given alone and in pairs. In this experiment, subjects sit next to each other in front of the machine. We did not record their discussion. Hence, the analysis is restricted to their answers, i.e. the options selected and the arguments introduced in the system to justify this choice. We describe these answers with three variables:

*3.3.1. Type of argument (solo and duo)*

- Accept: the argument provided supports the option selected, e.g. "I choose option 2 because it makes the learner autonomous".
- Discard: the argument rejects another option, e.g. "I choose option 2 because option 3 is too heavy to implement".
- Condition: the argument states some conditions for the option selected, e.g. "I choose option 2 if the learner is very young". Conditions can be imposed onto a certain category of learners (children, adolescents, adults, novices, experts...), or specific task domains (mathematics, verbal communication, procedural versus declarative knowledge...), or onto specific learning goals (consolidation, discovery...) or material settings (screen size, computer power...).

*3.3.2. Style of argument (solo and duo)*

- Justification: the argument contains at least one idea not present in the phrasing of the selected option
- Reformulation: the argument does not contain any idea different from those present in the selected option.

*3.3.3. Semantic relation between solo arguments and duo argument:*

- Union: the duo argument contains the ideas stated by both solo arguments
- Victory: the duo argument contains the ideas of one of the solo arguments
- New: the duo argument contains an idea, which is not present in either solo arguments

Table 1 illustrates the usage of this coding scheme for the solo and duo answers of one pair at question 1 (see figure 1 for the wording of the question). Subject1 solo argument shows a combination of several types of arguments: options 1 and 4 are discarded while option 3 is accepted conditionally. Notice that option 2, which is the actual choice of the student, is not addressed explicitly by the corresponding argument. Arguments with this level of complexity are quite rare, for most of the arguments were coded with a single type and style, like it is the case for subject2 solo argument. All three arguments were considered as 'Justifications' because they contain new ideas with regard to the option stated in the question. The duo argument is 'New' with respect to the content, because it introduces the idea of the learner's

## ELABORATING NEW ARGUMENTS

autonomy, which was not stated as such in the solo arguments. This examples illustrates what we call a conflict in this paper as the subjects gave different answers in the solo phase.

*Table 1. Coding example*

<i>Subject</i>	<i>Answer</i>	<i>Phase</i>	<i>Argument (we translate)</i>	<i>Coding</i>
1	Option 2	SOLO	If he is directly given the correct answer, he might not remember it, if we don't say anything, he might not be aware of his error... The help icon is certainly a good solution, if one can see clearly that it gives help and that it appears only on errors.	Discard option 1, justification Discard option 4, justification Condition for option 3, justification
2	Option 3	SOLO	trade-off between paternalism and neo liberalism	Accept option 2, justification
1 & 2	Option 3	DUO	It allows him to ask for help when he needs it.	Accept option 2, justification New Content

### *3.4. Results*

The system collected 180 solo answers and 90 duo answers to the questionnaire respectively. Some answers were incomplete, for instance when a student made a choice without providing an argument. The valid cases remaining for the analysis are 166 solo arguments and 67 duo arguments. The two authors of this chapter coded the arguments once and did not use a second judge to validate the coding. Yet, most of the coding was straightforward.

#### *3.4.1. The type of arguments varies between solo and duo phases*

Table 2 presents types of arguments for the solo and duo phases. We computed three proportion tests to compare the solo and duo phases. The proportion of 'Accept' arguments is similar in both phases even if significantly higher in the solo phase ( $U=0.048$ ,  $p<.05$ ). It appears that the proportion of 'Discard' arguments in the solo phase is higher than in the duo phase ( $U=0.037$ ,  $p<.05$ ). The proportion of 'Condition' arguments is higher in the duo phase than it is in the solo phase ( $U=0.034$ ,  $p<.05$ ). We will comment on the 'Condition' arguments when examining the effect of conflict on duo arguments.

Table 2. Type of arguments across solo and duo phases

		Phase				Difference
		SOLO		DUO		
		Count	%	Count	%	
Type of argument	Accept	136	82	51	76	- 6
	Condition	8	5	12	18	+ 13
	Discard	22	13	4	6	- 7
Total		166	100	67	100	

3.4.2. The style of arguments varies between solo and duo phases

Table 3 presents styles of arguments for the solo and duo phases. The proportion of ‘Reformulation’ is very high in the solo phase (more than half of the answers are simple reformulations of the choices proposed in the question). This proportion is much lower in the duo phase ( $U=0.06$ ,  $p<.05$ ). We interpret this difference as follows. When answering alone, students had no reason to make the effort to justify their choice with an elaborated argument even if they elaborated one in their head. In pairs, the discussion that was necessary to agree on an argument seems to force students to make it explicit and to elaborate it. This interesting observation confirms the usefulness of this method. The difference could however also be due to a simpler mechanism: the fact that the students have to answer twice to the same question might let them feel obliged to produce a more elaborated answer on the second time.

Table 3. Style of arguments across solo and duo phases

		Phase				Difference
		SOLO		DUO		
		Count	%	Count	%	
Style of argument	Justification	72	43	57	85	+ 42
	Reformulation	94	57	15	15	- 42
Total		166	100	67	100	

3.4.3. The type of duo arguments varies according to conflict

Let us now look at the duo arguments. Since we designed the learning activity in order to create conflicts, we will focus the analysis around this variable to see if it has an influence on the type and style of arguments.

Sometimes the AB pair chooses the answer given by A in the solo phase; sometime it chooses B's answer. We counted the pair symmetry, i.e. whether the pair chooses systematically the answers of the same peer. Data revealed that all pairs but

## ELABORATING NEW ARGUMENTS

2 were rather symmetrical, the difference between the number 'wins' of A and B ranging from zero to two (out of ten answers).

Table 4 presents the type of arguments for situations with and without conflict. In both cases 'Accept' arguments are predominant, students give an argument, which directly supports their common choice. 'Condition' arguments are interesting because they violate the instructions given to the students: they were asked to choose only one possible option and to give an argument supporting it. This way of proceeding would correspond to the production of 'Accept' arguments. Paradoxally, the 'Condition' arguments match the underlying pedagogical goals of the activity better than the 'Accept' type. We want to show the students that there is not one correct theory in courseware design but that each design choice depends on specific contextual features.

All but one 'Condition' arguments were given in a conflictual situation where students have to defend their choice against another's opinions. This type of argument is a way to solve the conflict by complexifying the argument. Indeed, establishing 'if..then..else' constructions allows to make concessions to the loser of the conflict. Unfortunately 2 cells from table 4 have expected count less than 5 so we cannot compute a reliable Chi-square test to validate these observations. Nevertheless, we modified the instructions given for the production of arguments and implemented a semi-structured interface to foster the production of 'Condition' arguments. The section about Experiment 2 reports the results we obtained.

*Table 4. Type of arguments across conflict*

		<i>Conflict</i>		<i>Total</i>
		<i>NO</i>	<i>YES</i>	
<i>Type of argument</i>	<i>Accept</i>	28	23	51
	<i>Condition</i>	1	11	12
	<i>Discard</i>	3	1	4
<i>Total</i>		32	35	67

### *3.4.4. The semantic relation between solo and duo arguments varies according to conflict*

The semantic relation between solo and duo arguments is related to conflict as can be seen in Table 5. When there is no conflict, students often take the arguments they gave in the solo phase and put them together to produce a 'Union' argument. Since there is no conflict, 'Victory' arguments correspond to the pair choosing one of the two arguments and use it as duo argument. When there is a conflict between the choices students made in the solo phase, the 'New' and 'Victory' arguments are more numerous. A 'Victory' in this context corresponds to one student winning the conflict on both aspects of the answer: the choice and the argument. A 'New' argument reflects a compromise, even if the pair chooses the choice of one student,

the argument they provide contains a new idea. 9 out of 12 ‘Condition’ arguments given in the duo phase are at the same time ‘New’ arguments.

Table 5. *Semantic relation between solo and duo arguments across conflict*

$(X^2 = 6.136, p = .047)$		Conflict		Total
		NO	YES	
Semantic relation	New	6	14	20
	Union	16	8	24
	Victory	10	13	23
Total		32	35	67

### 3.5. Discussion

Answering together to the questionnaire had an impact on the style of arguments (Reformulation vs. Justification) produced by the students. The duo arguments are more elaborated than the solo arguments. We propose that the reason for this change is that students make their opinions explicit and thereby more elaborated during the discussion. The type of arguments (Accept vs. Discard vs. Condition) also varies across solo and duo phases. ‘Discard’ arguments are more frequent in the solo phase while ‘Condition’ arguments are more frequent in the duo phase.

In the duo phase, half of the cases are conflictual, i.e. students did not make the same choice in the solo phase. ‘Condition’ arguments appear mostly in a conflictual situation. We have interpreted the production of this type of argument as resulting from a strategy intended to solve the conflict more fairly than by a ‘Victory’ of one of the students.

## 4. EXPERIMENT 2

### 4.1. Design and Research Question

Some students felt uncomfortable in having to make only one choice in experiment 1 and we noticed that forcing them to do so was in contradiction with the goal of the learning activity. With the version of the tool used in this experiment, students can accept or reject all alternatives by setting sliders to identical values.

The first experiment led to interesting results. Particularly, students accommodated to the conflict imposed to them by complexifying their arguments about design choices. This phenomenon met our pedagogical goals and thus, we wanted to further increase the production of conditional arguments. It led us to implement a modified answering modality that consists in rating possible choices on a five point likert scale. Labels for this scale were ‘Never’, ‘Rarely’, ‘Sometimes’, ‘Often’ and ‘Always’ (See figure 3).

## ELABORATING NEW ARGUMENTS

We were also interested in the argumentation going on during the duo phase. In order to record discussions we implemented a JAVA application that enables students to share a questionnaire and discuss through a chat tool while filling in the common choices and arguments. Whenever a student moves a slider or writes an argument into a text area, his or her partner immediately gets to see it. Having implemented that facility, it was possible to program the activity in a distance learning setting. Students were asked to do the solo phase at home during week 1. The system collected their choices and we run a pairing algorithm to match students with different opinions. Students then had to contact their designated partner and do the duo phase at their convenience during week 2. Finally, the debriefing took place during week 3 at the university.

**Enregistrer** **Next**

**Question 1 : Dans un didacticiel, si un élève commet une erreur, il est préférable de:**

1) Informer l'élève de son erreur et lui donner la réponse correcte  
Le plus rarement possible car l'élève peut faire moins d'efforts pour réfléchir

2) Informer l'élève de son erreur et lui fournir un indice qui l'oriente vers la bonne réponse  
Cela pousse l'élève à se débrouiller pour trouver lui-même la réponse.

3) Afficher une icône qui clignote, l'élève peut cliquer sur cette icône pour demander l'aide du tuteur  
Après quelques tentatives oui, mais pas à la première réponse fautive.

4) Laisser à l'élève un certain temps de réflexion pour qu'il trouve son erreur par lui-même  
Pour pousser l'élève, à au moins retrouver son erreur et réfléchir à la question, et proposer des solutions.

Figure 3. Question display in version 2. Students use the sliders to make a choice for each of the possible answers. They provide a written argument for each of their choices. The question and the options displayed are the same as the English version displayed in figure 1. The labels next to the sliders correspond to the scale ranging from 'Always' to 'Never'.

#### 4.2. *Subjects*

This experiment was run in 1999 between November 4th and November 30th. The participants were 16 students attending to the postgraduate diploma in educational technologies delivered by TECFA.

#### 4.3. *Variables*

Due to the new features of the tool, the variables used to code the arguments provided by the students differ from those used in experiment 1. Choices among alternative answers to the questions are not exclusive: it is possible to agree or to disagree with several alternatives by setting their respective sliders to the same value. The distinction between acceptance and rejection of an answer can be expressed by the manipulation of the sliders. Also, conflict is a continuous variable that depends on the difference of choices made by students.

##### 4.3.1. *Choice (solo and duo)*

The choice score reflects whether students select Never, Rarely, Sometimes, Often or Always. The distance between each label on this scale corresponds to one point. We use a numeric scale ranging from 0 to 4 to represent the selections of students.

##### 4.3.2. *Level of conflict (duo)*

The level of conflict for a particular alternative corresponds to the difference between the solo choice scores of the two learners. For a particular question, the level of conflict corresponds to the sum of differences over all the possible alternatives to that question.

##### 4.3.3. *Type of arguments (solo and duo)*

- Causal: the argument states a reason in favor or against the alternative. This category covers the Accept and Reject categories from experiment 1. Causal arguments are answers to the question 'why' and are represented by 'because' clauses.
- Condition: the argument states some conditions for the option selected by restricting its validity, e.g. "after some trials, yes, but not at the first error". Conditional arguments are answers to the question 'when?' and are represented by 'if' clauses.
- Other: there is no argument provided, the argument states subjective preferences for an alternative, e.g. "I like it better", non-arguments e.g. "I don't know, understand, ...".

Causal and Conditional categories might co-exist in a single argument. In those cases we assigned a Conditional code to the argument.

## ELABORATING NEW ARGUMENTS

### *4.3.4. Semantic relation between solo arguments and duo argument:*

- Union: the duo argument contains the ideas stated by both solo arguments
- Victory: the duo argument contains the ideas of only one of the solo arguments
- New: the duo argument contains an idea which is not present in either solo arguments

### *4.4. Results*

There were 26 alternative answers distributed over 9 questions. With 16 students this gives us 416 potential solo arguments and 208 potential duo arguments. Only one student out of 16 did not manage to fill out the solo questionnaire completely. She answered to the 3 first questions, totalising 9 choices out of 29 possible. Thus, the total number of solo answers collected by the system is 399. We had a big dropout rate in the duo session due to technical problems with the software and absenteeism. Three pairs did not manage to take the duo phase at all. Only five pairs remain in the analysis. Those pairs did not provide arguments for all the possible answers and skipped some questions. In the analyses concerning choices, we will use all the choices made by those five pairs, even if there was no argument provided. This corresponds to 260 solo arguments and 96 duo arguments. In the analyses concerning the arguments, we will eliminate the cases where there was no argument provided. This corresponds to 220 solo arguments and 88 duo arguments.

#### *4.4.1. The level of conflict is low*

The level of conflict across pairs is very low ( $M=1.04$ ,  $s=0.79$ ). Overall, on one particular duo choice, students were confronted with answers differing by only one point on the answer scale. The lowest mean of conflict for a pair was 0.6 and the highest was 1.4. These values are rather small and this is due to our pairing method, which gave the advantage to an even distribution of conflict over the intensity of the conflict. The distribution of level of conflicts is illustrated by figure 4.

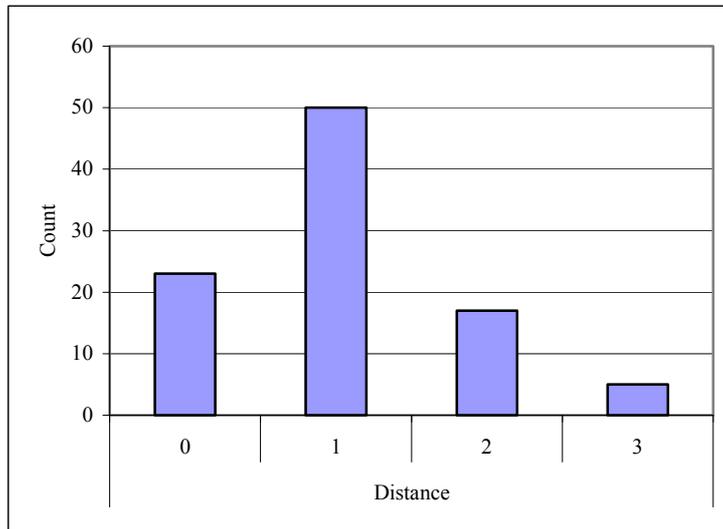


Figure 4. Distribution of Conflict levels ( $M=1.04$ ,  $s=0.79$ ).

4.4.2. The distribution of choices varies between solo and duo

The choices students made when evaluating alternatives range from ‘Never’ to ‘Always’ on a scale from 0 to 4. The mean value of choices in the solo phase is  $M=2.35$  ( $s= 0.98$ ). For the duo phase the mean is very similar ( $M=2.12$ ) but the standard deviation ( $s=0.48$ ) is half as big. This gives us a first indication about the choices in the duo phase, namely that the choices are centred around ‘Sometimes’ and that the extreme choices (‘Never’ and ‘Always’) are less present than in the solo phase. A test of variance confirms that the variance of the solo choices is bigger than the variances of the duo choices ( $F=4.47$ ,  $p<0.05$ ).

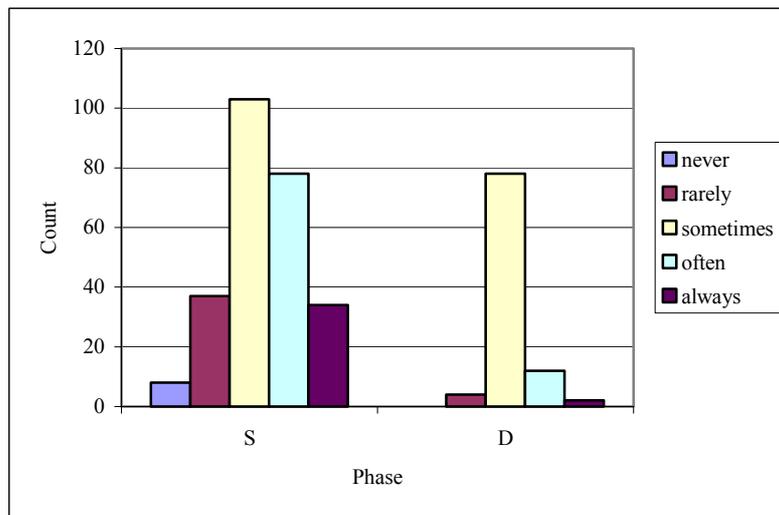
Table 6 shows that the proportion of ‘Sometimes’ doubled in the duo phase ( $U=6.36$ ,  $p<.05$ ) and that the proportion of ‘Always’ ( $U=.94$ ,  $p<0.05$ ), ‘Rarely’ ( $U=.87$ ,  $p<0.05$ ) and ‘Often’ ( $U=1.61$ ,  $p<.05$ ) decreased. ‘Never’ choices disappeared from the duo phase. A chi-square test confirms that Phase and Choice are dependant ( $\chi^2=49.897$ ,  $p=.000$ ). Figure 5 illustrates this change in the distribution of choices very clearly.

Table 6. Choice across solo and duo phases

$(\chi^2= 49.897, p = .000)$		Phase				Difference
		SOLO		DUO		
		Count	%	Count	%	%
Choice	Never	8	3	0	0	3

## ELABORATING NEW ARGUMENTS

	<i>Rarely</i>	37	14	4	4	10
	<i>Sometimes</i>	103	40	78	81	41
	<i>Often</i>	78	30	12	13	17
	<i>Always</i>	34	13	2	2	11
	<i>Total</i>	260	100	96	100	



*Figure 5. Distribution of choices in solo and duo phases*

At first glance, we interpret this move towards ‘Sometimes’ as a result of discussion taking place. Students would agree on choosing ‘Sometimes’ to handle conflictual situations. We will look later at the conflict variable to see if this is the case.

### *4.4.3. Type of arguments does not vary between solo and duo*

Student’s answers not only consist in setting the sliders but also in providing written arguments. Table 7 presents the type of arguments across the phases of the script. A chi-square test ( $\chi^2=2.246$ , NS) tells us that the type of argument is independent from the phase in which they were produced. When taking a closer look, it appears that the proportions of Causal, Condition and Other arguments do not change from the solo phase to the duo phase (all  $p>.05$ ). We expected an increase of ‘Condition’ arguments in the duo phase because in experiment 1 we considered them a privileged way to handle conflicts. However, in table 7, not all cases in the duo phase are conflictual and the increase of ‘Condition’ arguments might become apparent when opposing conflictual and non-conflictual situations.

Table 7. Type of argument across solo and duo phases

$(X^2 = 2.246, NS)$		Phase				Difference
		SOLO		DUO		
		Count	%	Count	%	%
Type of argument	Causal	161	73	62	70	11
	Condition	41	19	14	16	3
	Other	18	8	12	14	6
Total		220	100	88	100	

4.4.4. Choosing ‘Sometimes’ is not related to conflict

Do students use ‘Sometimes’ as a way to solve conflict? Examining the relationship between conflict level and choices will tell us more. As previously said, the level of conflict was rather low and because there are only few cases where the conflict level is greater than 1, we will aggregate the data and compare the cases where there is no conflict with the cases where there is conflict. In experiment 1, the definition of conflict was straightforward: students who chose a different alternative were in a conflictual situation. In this experiment, conflict is a continuous variable. Is a difference of 1 on the scale of choices perceived as conflictual by students? We propose two alternate definitions of conflict. We define weak conflict as follows: situations where the difference between students’ choices is equal to 0 are non-conflictual and situations where the difference is higher than 0 are conflictual. Strong conflict is defined as follows: situations where the difference between students’ choices is equal to 0 or 1 are non-conflictual and situations where the difference is higher than 1 are conflictual.

Table 8 shows the count and proportions of choices according to weak conflict. There are too many cells with expected frequency lower than 5 to compute a reliable chi-square test. The proportion of ‘Sometimes’ answers is not higher in the conflictual situations ( $p > .05$ ).

Table 8. Choice across weak conflict (levels 0 versus 1,2,3)

$(X^2 = 2.246, NS)$		Weak conflict				Difference
		NO		YES		
		Count	%	Count	%	%
Choice	Never	0	0	0	0	0
	Rarely	1	5	3	5	0
	Sometimes	15	71	55	82	11
	Often	5	23	7	10	13
	Always	0	0	2	3	3

ELABORATING NEW ARGUMENTS

	<i>Total</i>	21	100	67	100	
--	--------------	----	-----	----	-----	--

When using strong conflict we get the same configuration of proportions. Table 9 shows that the total of cases with and without conflict is the opposite with this definition of conflict, we have more non-conflictual cases than conflictual cases. Again, the number of cells with expected frequencies lower than 5 prevents us from using a chi-square statistic. When using a proportion tests, there is no difference between conflictual and non-conflictual situations regarding the ‘Sometimes’ choices ( $p > .05$ ).

Table 9. Choice across strong conflict (levels 0,1 versus 2, 3)

		<i>Strong conflict</i>				<i>Difference</i>
		<i>NO</i>		<i>YES</i>		
		<i>Count</i>	<i>%</i>	<i>Count</i>	<i>%</i>	
<i>Choice</i>	<i>Never</i>	0	0	0	0	0
	<i>Rarely</i>	3	4	1	5	1
	<i>Sometimes</i>	54	80	16	80	0
	<i>Often</i>	9	13	3	15	2
	<i>Always</i>	2	3	0	0	3
	<i>Total</i>	68	100	20	100	

We can now negatively answer to the question if the increase of ‘Sometimes’ choices in the duo phase results from the resolution of conflictual situations.

4.4.5. Type of arguments does not vary according to Conflict

Table 10 and table 11 show the type of argument across weak conflict and strong conflict respectively. In both cases, it is not possible to compute reliable chi-square statistics. Proportion tests comparing the frequency of ‘Condition’ arguments in conflictual and non-conflictual situations are not statistically significant ( $p > .05$ ). The difference of 12% for the proportion of ‘Condition’ arguments in table 11 is statistically not significant due to the small size of the sample. It nevertheless indicates a higher proportion of ‘Condition’ arguments in conflictual situations similarly to what was found in experiment 1.

Table 10. Type of argument across weak conflict (levels 0 versus 1, 2, 3)

		<i>Weak conflict</i>				<i>Difference</i>
		<i>NO</i>		<i>YES</i>		
		<i>Count</i>	<i>%</i>	<i>Count</i>	<i>%</i>	
<i>Type of</i>	<i>Causal</i>	16	76	46	69	7

P.JERMANN & P.DILLENBOURG

<i>argument</i>	<i>Condition</i>	4	19	10	15	4
	<i>Others</i>	1	5	11	16	11
	<i>Total</i>	21	100	67	100	

Table 11. s Type of argument across strong conflict (levels 0,1 versus 2, 3)

		<i>Strong conflict</i>				<i>Difference</i>
		<i>NO</i>		<i>YES</i>		
		<i>Count</i>	<i>%</i>	<i>Count</i>	<i>%</i>	<i>%</i>
<i>Type of argument</i>	<i>Causal</i>	53	78	9	45	32
	<i>Condition</i>	9	13	5	25	12
	<i>Others</i>	6	9	6	30	21
	<i>Total</i>	68	100	20	100	

4.4.6. More New arguments and less Union arguments are produced in conflictual situations.

Table 12 and table 13 present the semantic relation between solo and duo arguments according to weak and strong conflict respectively. With both weak and strong definitions of conflict, chi-square tests tell us that the semantic relation and conflict are dependant (weak conflict,  $X^2= 8.496$ ,  $p=.014$  and strong conflict,  $X^2= 6.558$ ,  $p=.038$ ). When looking closer at the proportions for the weak conflict we see a greater proportion of 'New' arguments ( $U=1.3674$ ,  $p<.05$ ) and a lower proportion of 'Union' arguments ( $U=1.5404$ ,  $p<.05$ ) in conflictual situations. With strong conflict, only the difference for 'New' arguments remains significant ( $U=2.0468$ ,  $p<.05$ ). There is no statistically significant difference concerning 'Victory' arguments with either definitions of conflict.

Table 12. Semantic relation across weak conflict (levels 0 versus 1, 2, 3)

		<i>Weak conflict</i>				<i>Difference</i>
		<i>NO</i>		<i>YES</i>		
		<i>Count</i>	<i>%</i>	<i>Count</i>	<i>%</i>	<i>%</i>
<i>Semantic relation</i>	<i>New</i>	5	24	35	52	26
	<i>Union</i>	10	48	12	18	30
	<i>Victory</i>	6	28	20	30	2
	<i>Total</i>	21	100	67	100	

## ELABORATING NEW ARGUMENTS

*Table 13. Semantic relation across strong conflict (levels 0, 1 versus 2, 3)*

		<i>Strong conflict</i>				<i>Diffenrece</i>
		<i>NO</i>		<i>YES</i>		
		<i>Count</i>	<i>%</i>	<i>Count</i>	<i>%</i>	<i>%</i>
<i>Semantic relation</i>	<i>New</i>	26	38	14	70	32
	<i>Union</i>	20	29	2	10	19
	<i>Victory</i>	22	22	4	20	2
	<i>Total</i>	68	100	20	100	

### *4.5. Discussion*

From the results we described, conflict seems to have had no statistically significant effect on the type of arguments provided by the students. We did not find evidence for an increased production of ‘Condition’ arguments in conflictual situations. Conflict had an effect though, as is shown by the increasing proportion of ‘New’ arguments and the decrease of ‘Union’ arguments in conflictual situations.

Few choices in the solo phase and almost no choices in the duo phase were labelled ‘Never’ and ‘Always’. Students didn’t take extreme positions, staying to the less implying ‘Rarely’, ‘Sometimes’ and ‘Often’. This seems to be a general phenomenon that is at work every time somebody has to answer to a Likert scale. Whatever the labels attached to the extremes, people tend to answer in the proximity of the centre of the scale. But here, this tendency to have moderate public opinions had a positive implication on the type of arguments. Half of the ‘Condition’ arguments were produced next to a ‘Sometimes’ choice.

## 5. GENERAL DISCUSSION

### *5.1.1. Conflict works*

Doise and Mugny (1981) proposed a theory of knowledge development rooted in an extension of the piagetian developmental psychology. They illustrated it by a series of experiments on socio-cognitive conflict. The basic idea of this theory is that social and cognitive factors are interdependent and both responsible for the development of knowledge. Also, there are two ways to handle a conflictual situation. A social resolution consists in conforming to the others’ opinion to reduce the discrepancy. A cognitive resolution consists in taking the position of the other party and adapting one’s own with respect to the other’s. The cognitive change then, results from the reduction of a social tension through perspective taking and accommodation of one’s own cognitive structures.

In our case, learning takes place as students evaluate alternative answers to a questionnaire. We want students to think about design choices by evaluating the

characteristics of the situation that present those choices. This means that we expect them to be able to justify any answer to the questionnaire without an important theoretical background. This approach works in this domain since questions are about educational choices and everybody has ideas about education; it would not apply to all domains.

When answering in pairs to the questionnaire, there were questions where students gave different answers and were de facto in a conflictual situation. The students use several strategies to handle this situation, each of which has implications for potential learning outcomes. The strategy closest to a social resolution of the conflict consists in taking both arguments and putting them side-by-side. There is no content-related discussion necessary to produce those type of answers. We described those cases as a 'Union' of solo arguments. Following a 'Victory' strategy, the pair takes one of the individual arguments and uses it for the pair. If the loser gives up without argumentation, this case corresponds to a social resolution as well. If the victory is the result of one student convincing the other and changing his beliefs, we are closer to a cognitive resolution. However, with this strategy, only one student changed his mind. Finally, producing a 'New' argument is closest to a cognitive resolution because the pair creates an argument that possibly matches both positions.

In both the first and the second experiment the proportion of 'New' arguments is higher in conflictual situations than it is in non-conflictual situations. The opposite is true concerning the 'Union' arguments; they are less frequent in conflictual situations than in non-conflictual situations. This result indicates that conflict is a useful mechanism to trigger deeper exploration of the domain. Students produce more new ideas when solving conflicting viewpoints.

Handling disagreement is a delicate matter especially when disagreement is imposed by an external means. As a matter of fact, our students did not choose to be in a conflictual situation. The resolution of conflicting opinions sometimes happens with a reference to justice. Students don't want to hurt each other and will make efforts to give the same amount of credit to each other. In some pairs' dialogues, the students manage to take answers from each other in an equilibrated manner. Other pairs, more comfortable with differences of opinions argue about the meaning they confer to concepts they used in the solo phase to find a common position. They also make the difference of opinions explicit in the discussion.

We should not over-generalize these results, which occurred on a very specific learning script. This script has been purposely designed for scaffolding rich interactions from conflicts. The necessity to be careful in generalizing these results is illustrated by their sensitivity to interface features, as explained in the next point.

#### *5.1.2. The characteristics of the tool matter*

Concerning the type of arguments, we were particularly interested in 'Condition' arguments. These arguments contain conditions and restrictions concerning the validity of a statement. For instance, a student might choose, "it is good to provide an animation feed-back on a correct answer" and, instead of providing an argument, adds the condition "IF the learner is a child". We saw that 'Condition' arguments

## ELABORATING NEW ARGUMENTS

appeared in experiment 1 mostly as a way to solve the opinion conflict. In experiment 2, the proportion of 'Condition' arguments was similar in the solo and the duo phase. And in the duo phase itself, there was no difference between conflictual and non-conflictual cases with this regard.

Even if the two experiments are presenting important differences, the questions and alternatives were the same. It is striking to see that in experiment 2 the conditional arguments are as numerous in the solo phase as they were in the experiment 1 duo phase.

We think that the sliders students use to make choices change the way they evaluate the options and provide arguments. For example, the labels 'Never' and 'Always' are adverbs with an absolute meaning. 'Never' and 'Always' imply a causal explanation, which we coded as either Reject or Accept in the first experiment and Causal in the second experiment. On the other hand, 'Sometimes' so to say 'calls' for a conditional argument stating the conditions in which the answer is true and the conditions that make it false.

For instance, it is grammatically more accurate to say "[Rarely or Sometimes] use video animations, they are distracting, except for young children" than to say "[Never] use video animations, they are distracting, except for young children". Never is never is... Even if less accurate from a grammatical point of view, arguments like the second example appeared in the data. Nevertheless, we think that the possibility to answer 'Sometimes' increases the number of Condition arguments as well in the solo as in the duo phase.

But why didn't the proportion of 'Condition' increase even more with conflict? An important difference between the experiments is that in the first one, students had to choose only one alternative while they had to evaluate all of them in the second experiment. In experiment 1, 'Condition' arguments also are a strategy to refer to other alternatives than the one chosen. Implicitly, saying A if X, can mean B if not X. Solving the conflict needed one student to change his or her choice. Considering that this was the only choice the pair could make, this change was an important concession. The pair would then set up an argument that compensates for the change made by one student by incorporating his or her previous choice as a 'Condition' in the new argument. In experiment 2, the production of several 'Causal' arguments can replace this need to refine the conditions for one choice This would explain why 'Condition' arguments do not serve conflict resolution in experiment 2.

### *5.1.3. The timing matters*

The two experiments also were differing with respect to the timing. Experiment 1 took place during a four-hour period with only 20 minutes separating the solo and the duo phases. During the debriefing, students were actively participating and justifying their arguments. Experiment 2 took place over a period of three weeks with one week between each main phase of the script. The participation during the debriefing was much less intensive than in experiment 1. The students didn't feel nor act accountable for the choices they made and the arguments they provided. Past

## P.JERMANN & P.DILLENBOURG

opinions were not strong enough to trigger another round of argumentation during the debriefing.

### 6. CONCLUSION

This chapter presents an analysis of arguments produced by students in a learning activity entitled “Argue Graph” where students twice answer to a questionnaire, a first time alone and then in pairs. We observed that answering in pairs had a positive impact on the elaboration of arguments provided to justify a choice in the questionnaire. We interpreted the improvement of arguments as stemming from the discussion necessary to give a common answer. Students make their opinions explicit and thereby more elaborated during the discussion. The tools’ features orient the way students perceive the task. The social implication of expressing opinions in these experiments may explain the differences in the type of arguments that were produced. We argue for a social psychological interpretation of the context of CSCL activities. Usually people tend to avoid social positioning by adopting non-implying positions. When a tool or a script allows them to stay ‘pat’ they will use this opportunity.

### 7. ACKNOWLEDGMENTS

TECFA Virtual Campus has been developed with the collaboration of Cyril Roiron, David Ott, Jean-Christophe Brouze, Daniel Schneider, Daniel Peraya, Patrick Mendelsohn, Philippe Lemay and Didier Strasser.

### 8. BIBLIOGRAPHY

- Blaye, A. (1988). Confrontation socio-cognitive et résolution de problème (A propos du produit de deux sous-ensembles). Doctorat de l'Université de Provence.
- Brown, A.L. & Palincsar, A.S. (1989). Guided cooperative learning and individual knowledge acquisition. In L.B. Resnick (Ed). *Knowing, learning and instruction, essays in honor of Robert Glaser*. Hillsdale, NJ: Lawrence Erlbaum Publisher.
- Butterworth, G. (1982). A brief account of the conflict between the individual & the social in models of cognitive growth. In G. Butterworth & P. Light (Eds) *Social Cognition* (pp. 3-16). Brighton, Sussex: Harvester Press.
- Chi M.T.H., Bassok, M., Lewis, M.W., Reimann, P. & Glaser, R. (1989). Self-Explanations: How Students Study and Use Examples in Learning to Solve Problems. *Cognitive Science*, 13,145-182.
- Doise, W. & Mugny, G. (1981). *Le développement social de l'intelligence*. Paris: InterEditions.
- Dillenbourg, P., Jermann, P., Schneider, D., Traum, D. and Buii, C. (1997). The design of MOO agents: Implications from a study on multi-modal collaborative problem solving. In *Proceedings of AI&ED'97 Conference*. B. du Boulay and R. Mizoguchi (Eds.). IOS Press.
- Dillenbourg, P. (2002) Over-scripting CSCL: The risks of blending collaborative learning with instructional design. In W. Jochens & P. A. Kirschner (Eds) *Three worlds of CSCL: Can we support CSCL?* Heerlen, The Netherlands: Open Universiteit Nederland.
- Hutchins, E. (1991). The social organization of distributed cognition. In L. B. Resnick , J. M. Levine & S.D. Teasley (Eds.), *Perspectives on Socially Shared Cognition*.
- Hutchins (1995). How a Cockpit Remembers Its Speeds. *Cognitive Science* 19, 265-288.
- Jermann, P. & Dillenbourg, P. (1999). Dialectics for collective activities: an approach to virtual campus design. In *Proceedings of AI&ED'99 Conference*.

## ELABORATING NEW ARGUMENTS

- Nickerson, R.S. (1993). On the distribution of cognition: some reflections. In G. Salomon (Ed) Distributed Cognition (pp. 229-261). Cambridge University Press, New York.
- Schoenfeld, A. H. (1987). What's All the Fuss About Metacognition ? In Cognitive Science and Mathematics Education. Ed: Alan H. Schoenfeld. London: Lawrence Erlbaum Associates.
- Smith, J.B. (1994). Collective Intelligence in Computer-Based Collaboration. Hillsdale, NJ: Lawrence Erlbaum.
- Schwarz, B. B., Neuman, Y., Gil, J. & Ilya, M. (2001) Effects of argumentative activities on collective and individual arguments. Proceedings of ECSCLE 2001, Maastricht, Netherlands.

*Patrick Jermann (Patrick.Jermann@tecfa.unige.ch)*  
*Pierre Dillenbourg (Pierre.Dillenbourg@tecfa.unige.ch)*  
*TECFA, University of Geneva; 40 Bd. du Pont d'Arve; CH-1203 Genève.*