Computer-Supported Collaborative Learning (2007) 2:297–313
DOI 10.1007/s11412-007-9020-6

Argumentation in a changing world

Baruch B. Schwarz · Reuma De Groot

Received: 14 April 2006 / Accepted: 16 July 2007 /
Published online: 6 September 2007

Abstract Critical reasoning has been recognized as a valuable educational goal since the end of the nineteenth century. However, the educational programs to reach this goal have changed dramatically during the twentieth century and moved to a dialogic approach. The shift to dialogism in programs to promote critical reasoning brings challenges concerning evaluation. We depict such a program here. This program is based on the use of graphic tools for argumentation in e-discussions. We focus on one history teacher who implemented the program in his class during a period of 7 months. In a design-based research cycle, we investigate the process of finding proper criteria to evaluate the program and to improve it. We show that the criteria of coherence, decisiveness and openness are appropriate for evaluating the program as they stem from pedagogical principles (autonomy, collaboration, commitment to reasoning, ethical communication, procedural mediation, etc.) that are central to a dialogic approach for critical reasoning education. We show that the history course was successful according to those criteria, but not successful according to other more traditional criteria. We discuss whether these differential performances suggest new standards for critical reasoning, actions to improve the program, or both.

Keywords Argumentation · Dialogism · Critical reasoning

New technologies and the promotion of critical reasoning: New hopes in the integration of dialogism in education

Tools are developed to facilitate human work. Well designed tools create a synergy with users by the means of which the functioning of the subject is modified (Heidegger 1996). However, the golem often rebels against its creator: The affordances and constraints induced by these tools engender new practices the designer did not even envisage that

To appear in the *International Journal of Computer-Supported Collaborative Learning*.

B. B. Schwarz (✉) · R. De Groot
School of Education, The Hebrew University, Mount Scopus, Jerusalem 91905, Israel
e-mail: msschwar@mscc.huji.ac.il

R. De Groot
e-mail: msruma@mscc.huji.ac.il

question the values, norms, etc., on which their design was rooted. Technological tools provide such kinds of surprises because their flexibility enables a very large range of behaviors. To cope with such unexpected outcomes, designers elaborate design cycles in which a new tool is first constructed; practices are then observed; researchers then reflect on the usability and efficiency of the tools, and on the values/norms challenged by the practices; and finally, the re-design of more appropriated tools is undertaken, leading to new observations, etc. The design process stands then in the middle of a societal activity in which human actions, tools and values are incessantly renegotiated. A well-known model for handling such issues is design-based research (Brown 1992; Cobb et al. 2001).

In the activity of education for thinking, many technological tools have been developed, and the Computer Supported Collaborative Learning (CSCL) community has been extremely active in this enterprise. The CSCL community has particularly been faithful to one of its theoretical objectives: reaching insights on learning processes involving the use of technological tools in rich social settings. CSILE was one of these tools (Scardamalia and Bereiter 1994). It conveyed the metaphor of "knowledge construction/building." This metaphor was used to observe how students could elaborate ideas during a short period of time (Bereiter 2002). The recognition that tools that encourage collaboration potentially foster cognitive gains led to the adoption of a participation metaphor of learning, a metaphor that had been expressed by Sfard (1998). Many of the tools elaborated in that second generation were argumentative representation tools (e.g., Belvedere; see Suthers and Hundhausen 2003; Suthers 2003). The representations used express a new direction in education for thinking that values (1) collaboration over individuation; (2) reasoning over thinking; and (3) reasoning activities over reasoning skills. Argumentative representation tools have recently been reviewed in several publications such as *Arguing to Learn* (Andriessen et al. 2003). Although many argumentative tools have been developed, results on educational programs for promoting reasoning with the help of argumentative representation tools during a long period of time are still missing. The program ITCOLE is an excellent example of such a program (see the special issue in *Computers and Education*, 2005, and especially Lakkala et al. 2005). However, although the program is aimed at supporting collaborative reasoning, collaboration is not really documented and evaluated (Schwarz 2005). In our view, this lack of empirical studies derives from several difficult issues. The first issue concerns the specific goals the educator intends to reach. The general objective of construction of knowledge or of participation needs to be translated into expectations in the specific context in which the program is implemented. A second connected issue concerns evaluation: What are the methodological tools that help measure the success of the program? A third issue concerns the consequences of the evaluation: If the program does not comply with our expectations, where should we seek for 'guilt'? Should we modify the program (tools, teachers, modes of intervention, activities, etc.), change the evaluation tools, or change our very expectations? Such questions characterize a societal shift concerning the value and definition of thinking and reasoning, and the cautiousness towards ideology that educators and researchers should possess when examining educational practices. This shift is deliberately towards reasoning as an argumentative activity in the framework of dialogism (Wegerif 2006). The philosophical and ideological ideas underlying this shift are nurtured by Habermas' (1972) Theory of Communication, Bakhtin's (1981) theory of dialogism, and Buber's (1923) philosophy of dialog. This shift motivated us to elaborate a new program, but we knew in advance that our expectations were vague, unarticulated and needed to be confronted with the practices developing with the designed tools.

Our goal in this study is to engage in the dangerous slope of design-based research, in which practices, norms and ideology are at stake and are adjusted through a design cycle.

The study stands in the middle of a design cycle with the DUNES environment http://dunes.gr), after the use of the tools was tested in controlled short term activities (Schwarz and Glassner 2007; Glassner and Schwarz 2005). We focus on a 7 months course implemented in a history class in which the teacher was dedicated to promoting critical reasoning according to the Kishurim program. We hypothesized that the use of CSCL tools should yield progress not only in promoting reasoning, but in instilling norms and values compatible with a philosophy of dialogue. To begin with, we describe the Kishurim general approach. This description clarifies what we mean by critical reasoning and by new norms and values compatible with a philosophy of dialogue. We list the pedagogical principles on which this program, dedicated to the promotion of critical reasoning according to dialogism, relies. We then describe the DUNES environment that provides tools, scripts, cases, etc. to enable the enactment of the principles.

Kishurim: A program dedicated to critical reasoning through dialogism

The Kishurim program is an educational initiative that both authors developed in Israel to foster argumentation and dialogic thinking in schools. The program has existed since 1998. The program includes pre- and in-service teacher's programs for helping teachers create animating activities in their classes. Although we developed several technology based environments to mediate argumentation and dialogic thinking, the program is dedicated to a pedagogy with and without technological tools. The Kishurim program does not provide predefined activities to be implemented, but leaves to teachers the autonomy to decide on the exact sequence of activities to be implemented in classes (Schwarz 2003; Schwarz and Glassner 2003). In order to promote this initiative, pre- and in-service teacher's programs are designed to accompany and support teachers in designing tasks according to their needs as well as in reflecting on them after their implementation in their classes. In these programs, researchers and designers negotiate with the teachers desirable practices in classrooms. We list here some principles. We stress that the listed principles are not ground rules to be blindly accepted by teachers, but are rather general recommendations for designing animating, and evaluating activities. These principles are used also to support ongoing discussions between teachers and researchers while working on the development of such activities.

Collaboration Tasks are often given to small groups in which participants are (made) aware that they contribute to a common goal and that their collaboration is in itself valuable.

Non-intrusive procedural mediation Teachers are suggested to use procedures that support collaboration (in electronic and face-to-face modes) to promote critical reasoning: Teachers are invited to decide on specific scripts, general instructions given beforehand and possibly during their activity. Based on research, we insisted that inviting students simply to discuss an issue often does not promote rich discussion and that a guideline such as "try to reach common understanding" or "try to accommodate your different opinions" were more useful. Another procedural mediation consists of using ontologies provided by the representational argumentative tool. Teachers are invited to minimize use of authority and direct interaction with students.

Commitment to critical dialogue This is one of the central principles in the program. Teachers are committed to inviting students to give reasoned arguments, to open new perspectives, to challenge arguments on which they disagree, and to negotiate/revise their

arguments when data or incompatible arguments cannot be rejected. The main tool utilized is argumentation, dialogic patterns through which construction of knowledge may be reached.

Ethical communication Communication, especially in critical dialogues, raises ethical issues; for example, concerning how to account for flawed reasoning, whether to favor arguments in light of their quality without taking into account the identity of their authors, etc. This aspect is particularly complicated to handle when participants are committed to acquiring scientific knowledge that generally cannot be put into question (and if it is challenged, students and teachers know that it will only be established to conform to the norms of the domain).

Autonomy In parallel with the ethical aspect of communication that is handled by appropriating ideas to the group, the teacher aims at giving the opportunity for each student to develop ideas by his/her own, in his/her interactions with peers. As most classes are cognitively heterogeneous, since we also want to preserve ethical aspects of communication, autonomy cannot be fostered unless the teacher interacts differently with different students.

Active role of the teacher in the design of activities The role of the teacher is especially important in deciding in advance on the kind of dialogue in which to engage. With the DUNES environment, it means in particular deciding on an appropriate ontology (shapes in the Digalo tool that point at categories of talk), on the script to prepare for e-discussions, and on the kinds of interventions to be prepared to implement. Design also means transforming activities generally administered in a traditional didactic manner into argumentative dialectic activities.

Disposition to exploit resources to instigate dialogue This principle expresses a positive attitude toward the use of any resource or tool that can enrich classroom activity. As for resources, this means that the teacher is committed to providing texts, data, etc. that can foster reasoning. As for tools, we elaborated the DUNES environment, which provides a platform for e-discussions in synchronous or asynchronous mode. DUNES has been an EC funded project (IST-2001-34153) coordinated by the authors of this paper. The objective in DUNES was to design, implement and test an environment for collective argumentation. The main technological outcome of this project was the development of Digalo, a graphical tool representing synchronous and asynchronous discussions. In order to motivate students to engage in discussions, we followed the advice of several researchers (van Bruggen and Kirschner 2003; Schwarz and Glassner 2003) to propose to students 'cases' (based on 'ill-structured' or 'wicked problems'); that is, problems for which (a) there is no unique expected answer, (b) the ways to progress to an acceptable solution are varied and (c) participants have some informal knowledge. Teachers design cases and decide on the social settings of activities, such as the size of groups of discussants (small groups of two to six, whole group forum, or individuals). The script of a case contains definitions for the schedule, pedagogical goals, content-related goals, etc. An example of a content-related goal is to differentiate between the role of primary and secondary texts in the elaboration of interpretations in history issues. The pedagogical goals are generally non-content-related goals that teachers expect the users to learn and express the dialogical principles on which we intended to rely. For example, goals such as learning how to reach a better common understanding or how we accommodate divergent views express principles of collaboration

and of commitment to critical reasoning. A goal such as learning how we trust (or not) what somebody is telling us expresses the principle of ethical communication. These pedagogical goals are not spelled out explicitly, but expressed through prompts that the teacher invites students to follow. For example, when the teacher aims at reaching common understanding, prompts such as "try to refer to each other," or "do you agree with" can be used. In activities in which the aim is to confront and accommodate divergent views, prompts could be "try to challenge this view" or "are you sure that your conclusion necessarily follows the data you have." Such prompts, which express the principle of procedural mediation, are intended to promote critical dialogues in which discussants gain autonomy.

Our experience in animating argumentative activities in classrooms led us to develop the Digalo tool http://zeno8.ais.fraunhofer.de/digalo/index.html) within the DUNES project. This tool enables the management of argumentative discussions and the representation of argumentative processes and components among participants. Digalo consists of co-created maps built of written notes inside different shapes. Different arrows (supporting, opposing, and linking) represent different connections between the shapes and collective argumentation enables reference to each other's ideas. Every map has an ontology that specifies and constrains not only the admissible labels for the shapes (such as opinion, fact, reason, defending, challenging), but also the different 'roles' to be played when manipulating the map (the shapes can be configured according to the teacher's decision). The choice of ontology is intended to create a discussion space to promote coherence and connectivity between the participants. When using Digalo, the facilitator of the discussion (generally the teacher, but in some cases a student) presents a blank map and decides on the ontology to be used in e-discussions. The modes of communication in Digalo may be verbal or electronic in synchronous or asynchronous discussions. Verbal and electronic modes of communication may concur when students are in the same computer room. The remote mode of communication may include the synchronous use of the Pad in Digalo with a chat channel.

Figure 1 shows a part of an e-discussion between students in a class. We will discuss it later on. One may already see that the ontology includes "claim," "argument," "information," and "question," and the arrows "support," "oppose" and "link." Each discussant has an identification badge, an important fact in e-discussions. Figure 1 shows the contribution of the student Shir, which is a claim, whose title is "children and parents" and whose content shown in the lower text box begins with "I think that children were influenced." The colored boxes are the teacher's interventions.

A design-based research program was launched in which many teachers committed themselves to critical reasoning through dialogism. One of the teachers taught a 7-month course on a historical theme. This course served the design-based research team evaluating the success of Kishurim and reflecting on the adequacy of the very tools used for this evaluation.

Description of the experiment

The history course

The experiment took place in the framework of a history course. The course took place in an Israeli urban, non-selective, academically oriented high school. The principal was interested in promoting critical reasoning in classes. Her overall objective was also to

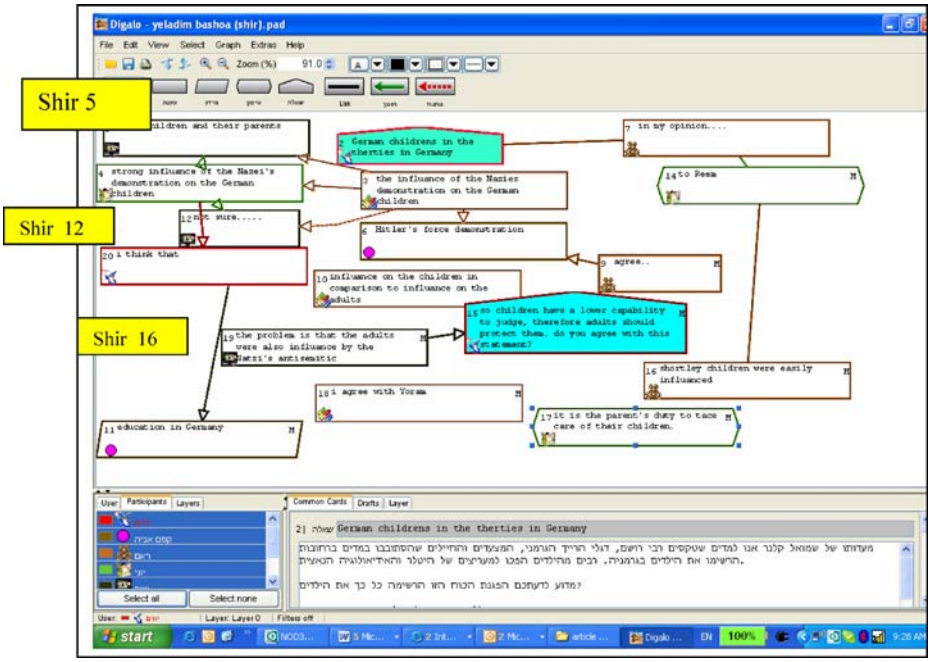

Fig. 1 Part of an e-discussion between students in a class (contribution of the student Shir)

encourage the use of new technologies in learning activities and to encourage teachers to initiate autonomous extracurricular activities in order to promote leadership in the teaching staff of the school. As a part of an in-service teachers' program, a history teacher designed a series of activities aimed at developing ideas about the theme "childhood and war." The teacher had taught for 4 years; he was used to the integration of various technologies in his courses (history database for preparing lessons and animating them, use of portfolios to follow the writing of essays in history, etc.). The teacher had participated a year before in an in-service teachers' program in which the principles were presented to teachers and negotiated, then implemented in the teachers' classrooms and reflected on in further sessions with the teachers. Ten Grade 9 students (five males, five females) from the same class participated in the course. Students were free to choose one among all such activities, and the ten students decided to take part in the course on "childhood and war." The teacher focused more on Jewish children during World War 2, but general issues around different wars and about children from different countries in different times were raised during the program. The design of the activities generally seemed to encourage the principles listed above. For example, a lot of activities necessitated collaboration; critical reasoning was favored in collective e-argumentation; and the teacher intensively exploited textual resources.

The course began with a review of crucial events in World War 2, on the ghetto policy, and on some data on children during the holocaust. Then students read various sources, such as children's diaries or texts written by adults on children during the holocaust. While or following the reading, children asked questions, participated in discussions, and expressed their opinions. At the end of the course, students submitted a memoir in which they were asked to present a reasoned and elaborated argument about childhood and war.

Intertwined with the history activities, students were taught how to formulate an inquiry question, how to describe the inquiry process in a structured way, and to express the

conclusion. They were instructed in argumentation: claims, arguments, and reasons were defined. They were also given criteria according to which arguments could be evaluated and were also instructed to use a given ontology in Digalo to construct argumentative maps. While teaching students the use of this ontology, the teacher formulated ground rules about dialogue. These rules were almost always presented in the context of history. For example, when encouraging students to explain their claims, the teacher often added remarks such as "in history you can't write down your opinion without relying on explanations or facts" or the like. When he introduced arrows of opposition, the teacher explained that in history, one should always think about alternatives to reach the truth. The teacher also used such remarks in activities in which students worked in small groups to elaborate interpretations or in plenary discussions as moderating actions. The course lasted 7 months. It was held in 24 weekly 90-min long meetings in the computer room at school, or in virtual meetings as students remained at home. All tasks were posted in a portal in the DUNES environment. In some cases, these tasks served for the preparation of the Digalo session to be held in the classroom. All lessons were videotaped. In all activities, participants used the Hebrew language.

Research questions

Two conflicting research questions were pursued in parallel. The first one concerned the success of the program and asked whether the students improved their historical reasoning on the issue at hand, childhood and war, during the 7 months of the experiment. The second question was methodological: we asked what the appropriate tools are to measure the success of the Kishurim program. A priori, the juxtaposition of the two research questions is not legitimate: by asking the second question, we question the appropriateness of the tools used to measure the success of the program. However, given the embryonic state of methodology for measuring the efficacy of programs intended to promote critical reasoning through dialogism, our approach means that for the first research question a variety of measures is used, and the second question is a reflection on whether the measures really reflect success or failure from an educational point of view, or rather that some of the measures used are inadequate.

Tools for undertaking evaluation

It was clear to us that measuring reasoning in history did not concern facts, but the way they organize their arguments and interpretations on historical events (Wineburg 1991). Students were invited to write a short essay to evaluate their reasoning about childhood and war. Students were asked "what kind of childhood did children experienced during the holocaust? Give as many different reasons as possible". In addition, we collected all the Digalo maps produced during the experiment. The rationale for writing the essay was that since the teacher always aimed at promoting reasoning in history, comparison between essays that demanded the elaboration of interpretations would uncover understandings developed during the activities. We knew, of course, that the writing of the essays did not involve an overt dialogue with others, and as such would risk not bringing to the surface dialogic aspects in which we were interested. However, we hypothesized that identifying argumentative components of the essays would reflect some footprints of the dialogical activities. Our hypothesis was not precise as we did not know for which argumentative component effects would be identified. We decided that direct analysis of dialogic activities would be a second step only in the evaluation, especially since the evaluation of tools for

direct evaluation of dialogic activities is still in an embryonic stage. In the present paper, the direct evaluation of dialogic activities is proposed to interpret findings in the essays.

Analysis of the data

In the essays, we identified the number of claims, reasons supporting claims, the openness, the coherence, the change of opinion, and the decisiveness of the claims. The number of claims and reasons supporting the claims are traditionally used for evaluating reasoning skills (Means and Voss 1996; Kuhn 1991) and for measuring knowledge construction and reasoning (Schwarz et al. 2003). The other measures emerged as the tests were examined and the researchers judged them as relevant to the study (see the next section). All criteria are clarified in an example, the initial and final essays that Nathalie, one of the students, wrote. Squared brackets point out at claims and reasons.

Initial essay: "In my opinion, the kind of childhood children had during the holocaust is not defined [claim 1], that is to say that the kind [of childhood] depends on the place and the time, and on the child. In general, the childhood, in my opinion, ended very quickly [claim 2]. This was a reality that imposed a quick maturation from a mental perspective [elaboration of claim 2], since everything was done fast, without clear answers [reason 1 to claim 2]. Generally, children expect from adults—parents, neighbors, relatives, to explain and to answer their questions. At that period, nothing like that happened [reason 2 to claim 2]. Childhood was always accompanied with constant fear, lack of knowledge concerning the future [claim 3]. Shall we live? Shall we see our family again? How much time we'll stay alive? Why should we die? [elaboration of claim 3]."

Final essay: "The children that were in the holocaust experienced a kind of childhood that was introverted [claim 1]. The new framework in which they were forced to live influenced very much their way of life and their daily behavior [claim 2]. They were forced to take care of themselves by their own [reason 1 to claim 2]. The activity needed to do that, took most of their time and necessitated a serious and mature attitude [reason 2 to claim 2]. Childhood was always dormant but burst only from time to time [reason 1 to claim 1]. One should notice that the situations in which they were hurled influence every child differently."

Three claims were identified in the initial and the final essays. Two reasons were invoked in the initial essay, and three in the final essay. Openness was measured through the number of perspectives invoked. Two perspectives were adopted in the two tests: a pragmatic perspective ("will we live?", "they were forced to take care of themselves," etc.) and a psychological one ("This was a reality that imposed a quick maturation from a mental perspective," "The children that were in the holocaust experienced a kind of childhood that was introverted"). The coherence was measured through connectedness, a link between ideas presented, the link being logical or chronological. The score for coherence ranged from 1 (without coherence), to 3 (full coherence). A score of 2 was given for partial coherence. In Nathalie's initial essay, coherence was only partial, as the link between claims 1 and 2 is not clear. In the final essay, coherence was partial too, since the reason to claim 1 was disconnected from claim 1. The arguments developed in the initial essay—the fact that there was no real childhood during the holocaust—and the final essay—the fact that childhood was introverted—show a change of opinion. As for decisiveness, it conveys the determination and the principled way expressed in the argument chosen. In the initial essay, Nathalie expresses doubts and uncertainties ("the kind of childhood children had during the holocaust is not defined") in the argument as well as specific questions ("Shall we live? Shall we see our family again? How much time we'll stay alive?"). In contrast, in the final essay, the argument is expressed with determination in a principled way ("The children...

experienced a kind of childhood that was introverted"). Especially for coherence and decisiveness, scoring resulted from a subjective evaluation. Consequently, three judges analyzed the initial and final essays independently and disagreements were negotiated. The Cohen's Kappa test yielded a measure of reliability of more than 0.80 for each of the variables.

Results

The results are displayed in Table 1. The comparison between the essays written in the pre and post-tests did not show differences for the variables that traditionally measure the (improvement of) argumentative components (number of claims, explanations and reasons). Improvement could be found for openness (the number of perspectives increased from 16 to 22, p=0.012), the coherence of texts (from 23 to 27, p=0.052, almost significant) and for decisiveness (from 6 to 14, p=0.026). As for the change of opinion, six students changed their mind (change is marked as '1' and stability '0' in Table 1). Although we presented openness, decisiveness, and coherence as dependent variables in the experiment section, we admit that we did not think about them in advance. The analysis itself provided ideas to define new directions. These new directions precisely express the gist of the design-based research approach we adopted to evaluate the effects of the Kishurim program. It also answers to the antagonist research questions we already asked: According to criteria that 'traditionally' measured knowledge construction (number of claims and reasons), the experiment failed to reject the null hypothesis. But for the second research question concerning the adequacy of the methodological tools chosen to measure the impact of the Kishurim program, the answer is negative: other tools are needed. We propose criteria such as openness, decisiveness and coherence. But to definitively adopt these new criteria, we need to identify new practices and outcomes that may explain our findings and to show that these criteria fit the pedagogical principles of the program. For this purpose, we decided to observe activities that may explain some of the changes in Table 1. We present one such activity here.

Table 1 Results

Name	Claims		Expl.		Reas.		Pers.		Cohe.		Cons.	Decis.	
Nathalie	3	2	4	1	2	3	2	2	3	3	1	0	1
Kesem	1	1	0	0	5	2	2	2	1	2	0	0	1
Ben	2	3	0	0	3	2	1	2	3	3	1	3	2
Reem	2	3	0	1	3	2	2	2	2	3	1	0	2
Yoni	3	3	0	0	4	3	2	3	3	3	0	1	0
Ran	3	1	1	0	1	2	2	2	2	3	1	0	2
Shir	1	3	1	1	0	0	1	3	2	3	1	0	1
Gal	2	3	1	0	2	2	2	3	1	2	1	1	1
Keren	1	3	1	1	1	4	1	2	3	3	0	1	3
Linoy	1	1	1	1	5	3	1	1	3	2	0	0	1
p values		0.172		0.135		0.296		0.012		0.052			0.026
Mean	1.9	2.3	0.9	0.5	2.6	2.3	1.6	2.2	2.3	2.7	0.6	0.6	1.4
SD	0.876	0.949	1.197	0.527	1.713	1.059	0.516	0.632	0.823	0.483		0.966	0.843

Analysis of dialogues in search of improvements in openness, decisiveness and coherence

Many discussions took place during the history course. Ten of them were mediated by the Digalo tool. Fourteen discussions were done in a face-to-face mode. In only two of the ten Digalo-mediated activities, could students fully express their viewpoints in an autonomous way, compare them with other viewpoints and texts and subsequently revise them. In other activities, the argumentative activity was scaffolded by the teacher through questioning techniques, the students being often in the position of reactants. We hypothesized that the two activities in which more autonomy would be given to students would provide more information about the views they developed on childhood and war. The activity we chose to focus on, "Children in Germany in the 1930s," is one of these two activities. This activity is typical in the sense that the teacher reads with the students a historical source, asks for their personal opinion, and invites them to participate in a discussion in order to answer a general question he asked. The discussion we describe took place in the computer lab. Students used Digalo. The discussion map they produced is displayed in Fig. 1. We focus on how Shir, a girl in the class, participated in the discussion. She belongs to a group of four children (Shir, Ben, Yoni, Kesem). Other interventions are only hinted at for reasons of place. Our goal is to illustrate that openness, decisiveness and coherence, whose improvements were identified in the post-test essays, can be explained in the dialogue and that this explanation accounts for the principles on which Kishurim was based.

Typically, the teacher has first prepared a text in order to trigger the discussion. As usual, he was *actively engaged in designing this activity*, a central principle in Kishurim, and *exploited textual resources to instigate dialogue*. After reading the text, Yoram asks a question (in the Digalo map):

> Yoram: Title: Children in Germany in the thirties. Comment: From Shmuel Kelner's testimony, we learn that impressive ceremonies, Reich's flags, marches, and soldiers in the streets did fascinate children in Germany. Many kids turned into admirers of Hitler and of the Nazis ideology.
>
> In your opinion why did this demonstration of force fascinate the children so much?
>
> Do you think that children were more influenced by these demonstrations than the adults? What can you learn from it?

This question was posted beforehand on the Digalo screens of one group of students. At that stage, the students are used to expressing first their personal opinion. This is a clear invitation to *autonomy*. Yoni writes the following contribution:

> Yoni (3): Title: The strong influence of demonstration of power on German children. Comment: In my opinion the demonstrations of power of the Nazis influenced children in Europe because children admire strong personalities (the Nazis invaded vast parts of Europe and were a strong force in Europe). For example, young children admire Superman and Spiderman, because they are strong. At that time the Nazis showed that they were very strong and children were influenced. The demonstrations had more impact on children as I explained it above.

We underlined the last sentence to express that Yoni wrote his first contribution and posted it without the last sentence before Shir. Shir then saw what Yoni had posted and then

posted her own comment. Yoni added his last sentence after he saw Shir's contribution. Shir's first contribution is the following:

> Shir (4): Title: children and their parents. Comment: I think that the children were highly influenced by the fuss the demonstrations did. Such big dimensions show tremendous power to which kids cannot resist. As opposed to children, their parents could demarcate themselves from these demonstrations without being carried away by anti-Semitism. Unfortunately part of them [the parents] was influenced by these demonstrations so they collaborated.

We can understand why Yoni added his last sentence to the opinion he expressed before. He probably totally agrees with the idea expressed by Shir that the parents could resist the Nazi propaganda but the children could not. He probably added *The demonstrations had more impact on children as I explained it above* to appropriate Shir's idea without overtly admitting it. The synchrony and the fact that students could modify their past contributions enabled students to pick up ideas from others without losing face. Consequently, although students worked together, Digalo enabled autonomy in discussions.

After students expressed their personal opinions, they engaged in collective argumentation. At that stage of the course, collective argumentation was not triggered by the teacher (see also van Diggelen et al. 2005). Rather, it turned to a norm:

> ...Shir (10): Answer to Kesem at turn 9 (Kesem said that the children were influenced more than their parents and that they convinced their parents) Title: not sure Comment: You think that the children were more influenced than their parents? In my view this is not true; the fact is that many adults supported the Nazi ideology because of these demonstrations too.

This reference to others characterized collective argumentation and expressed a *collaborative norm* (another principle in Kishurim) even when, such as in this case, discussants did not agree. As in the beginning of the dialogue, when students spelled out their personal opinions, collective argumentation uncovered reasoned arguments, challenges, etc. that showed *commitment to critical reasoning*. The teacher attempted *to mediate the promotion of critical reasoning;* at turn 13, the teacher intervenes by summarizing Shir's intervention and by asking the group whether they agree:

> Yoram (13): so the kids have a lower ability to judge the situation so the adult's role is to govern them. Do you agree with such a statement?

Shir reacts in Turn 16:

> Shir (16) Title: the problem is that also the adults were flooded by the anti-Semitic wave Comment: Indeed the role of the adults is to take over their children and teach them not to believe to anything they see, but in the case of so many demonstrations and the activity of the Nazi ideology, the adults were also deceived by the credibility of the Nazi party.

Overall, this example is interesting from several aspects. First, the thread of the discussion is quite long: there are 16 turns until Shir expresses her "final" opinion. Second, Shir's interventions refer to previous interventions of their peers: Yoram's question in turn 4, Kesem's claim in turn 10, and Yoram's question in turn 16. As we can see it, her interventions do not follow immediately the interventions to which they refer. Third, many interventions are reasoned. Also, the discussion is symmetric in the sense that it is equally

distributed among participants, including the teacher (this cannot be fully illustrated here due to space constraints; Fig. 1 suggests this symmetry). All aspects generally do not characterize school discussions. Since, as shown by Cazden (1988), classroom discourse generally appears to be *asymmetric*, with a tripartite initiation, response, evaluation (IRE) structure with short threads in classroom discourse, references tend to follow immediately a question (generally by the teacher), and students often "agree" with previous claims proposed by the teacher. In the present discussion, on the contrary, students are at the center: Shir begins with the description of a situation in turn 4, continues with the reasoned reference to another view in turn 10, and a conclusion that takes into consideration a question by the teacher, which itself took into consideration different opinions. Shir's opinion making results from a rich social interactive process in which Digalo plays an important role. The persistence of previous interventions and the possibility to privately edit interventions while looking at the growing map helps students to make up their opinions autonomously, but through collaboration with others. We suggest that such Digalo affordances can explain the increase of decisiveness. Yoram mediates critical reasoning processes in a non-intrusive way: he often summarizes previous turns (like *so the kids have a lower ability to judge the situation so the adult's role is to govern them* in turn 16) and asks a question (like *Do you agree with such a statement?* at the end of turn 12) to let students develop their own ideas. Consequently, the length of threads and the references that students make to previous interventions suggests the instillation of the combination of principles on which Kishurim relies, especially *commitment to critical reasoning* and *non-intrusive procedural mediation*. We suggest that this combination naturally leads to high coherence in collective argumentation and that it is internalized in post-test essays. The interventions by Reem, another student, in the same group (see Fig. 1) confirm this suggestion. At turn 5, he expresses his personal opinion:

> Reem (5). Title: In my opinion. Comment: All the demonstrations and the soldiers in the streets influenced the children because the children saw something bigger than them and believed that the adults are doing something big and important!

Reem engages then in collective argumentation:

> Reem (7). Title: I agree. Comment: I totally agree (with Kesem in 6). In my opinion children are more easily influenced by Hitler's power demonstrations because children seek for a figure for imitation and at that period, Hitler presented himself as a powerful figure, so that children saw in him the image of a hero.

In turn 13, she expresses her conclusion:

> Reem (13): In short, children were easily influenced.

This short list of interventions shows the same phenomena we saw with Shir: Reem expresses a reasoned claim in turn 5 and agrees with another reasoned claim in turn 7. In that turn, his position is expressed with a lot of determination ("I totally agree with," "in my opinion"). It is summarized in turn 13. Like Shir, Reem expresses herself as equal in a reasoned discussion where the teacher elicits 'argumentative skills,' but is not intrusive.

We present now another e-discussion with another group that produced the argumentative map displayed in Fig. 2. Similar to the former one, this discussion follows the reading of a text and begins with a question posed by the teacher:

> Yoram: Why they did not escape from Germany timely?

> Dov Amir argues in his testimony that the parents had routine habits of mind that avoided from them to understand the reality and to anticipate future dangers. It seems

that Dov Amir is angry at the silliness of the parents. Please try to enter the head of Dov Amir's parents in the Nazi period in the thirties and write down what prevented them from deciding to escape from Germany after Hitler took over the country?

Once again, the discussion begins by the expression of personal opinions. Here again, we focus on one student, Linoy, who belongs to a group of four. When Linoy expresses her opinion, she expresses it as an opposition materialized by an arrow to an opinion expressed in the text (presented by the teacher) in which it is claimed that "parents were silly." She uses the form of an argument to express her intervention (see Fig. 2):

Linoy (6): Title: It's not simple. Comment: It's not simple to leave the country in which you lived so many years. Therefore, it's difficult to take the decision to leave your home.

As collective argumentation develops, Linoy requests further information:

Linoy (11): Title: I have a question. Comment: What was the situation of Germany in 1939 and before?

Here, Linoy uses the appropriate form: a "question" form. Skipping to turn 15:

Linoy (15): Title: Still, it's not so simple. Comment: To leave your home, whatever bad the situation is, is not simple and easy. No matter the kind of hell you are in. But the truth is that, after all, one should take the right decision, even if this is very tough!

In turn 20, as the discussion develops, Linoy requests assertively an answer to her unanswered question:

Linoy (20). Title: Still you did not answer my question! Comment: My question was about the security situation before 1939

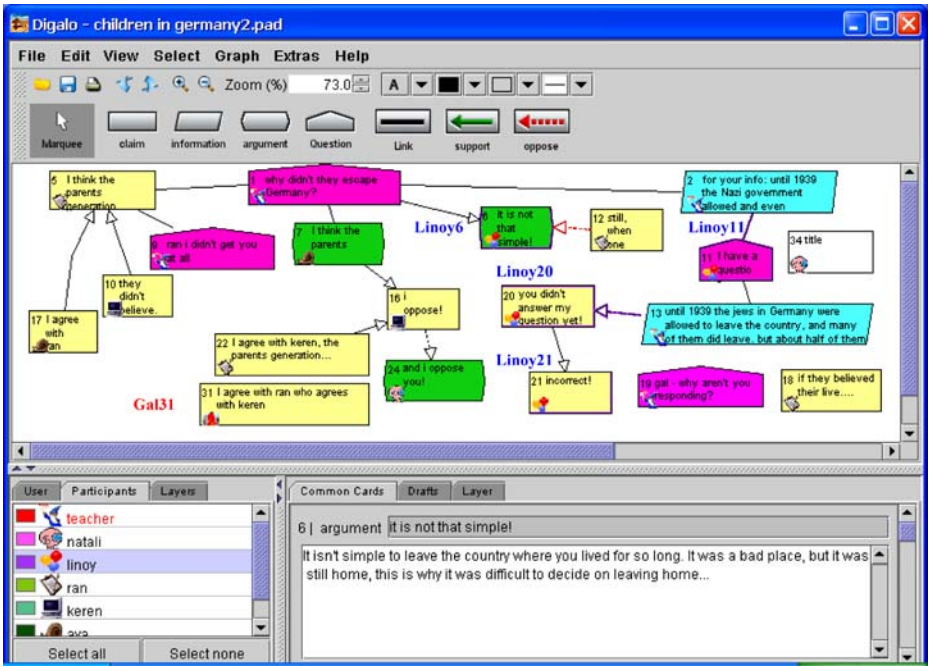

Fig. 2 Argumentative map of a second e-discussion

In turn 21, she summarizes her conclusion as an opposition:

> Linoy (21). Title: Not correct! Comment: In my opinion, they believed that their lives were endangered but they used their hearts and not their heads that told them to escape from Germany when this was possible

In turn 6, Linoy first shows empathy to people who did not escape in time. She is not sure about her own position ("it's not so simple"). In turn 11, she initiates the checking of historical data. In turn 15, she reiterates her empathy. In turn 20, she reiterates her request for data, and in turn 21, she expresses her conclusion with determination. Similar to the first example, the thread of discussion is long, involves interaction with others, and is highly reasoned.

The examples we presented show recurrently that the teacher structured the activity very well by reading a historical text, asking for personal opinions and summarizing the opinions of students. He could structure the discussion in such a way since he could write questions in advance in Digalo's discussion space and could capitalize on the persistence of contributions of students in the map to ask questions or summarize. He let students spontaneously engaging in collective argumentation. This engagement was probably influenced by the personalization of interventions by icons that functioned as representational affordances in Digalo. We suggest that the structure instilled by the teacher and the collective argumentation can explain the coherence of ideas that developed, one of the unexpected gains detected in the post-test. Also, the diversity of perspectives that expressed openness was embedded in the structure of the discussions we presented. As for decisiveness, it is well rendered by the equal distribution of contributions, in the request for personal opinions, and in the importance given by the teacher for these opinions in collective argumentation. Also, the use of testimonies in history classes encourages students to function as actors in the construction of interpretations (see Linor in turn 11 and her insistence in turn 20).

Discussion

The issue we discussed in the argumentative activity undertaken with the help of Digalo concerns the evaluation of argumentation products and processes. We adopted a "successive approximations" approach: we first used a pre-test/post-test paradigm, through which we evaluated individual written essays before and after the series of activities in the Kishurim program. The criteria for evaluation concerned argumentative characteristics of texts (number of arguments, number of reasons, etc.). These criteria had already been used by several researchers to measure the acquisition or the development of argumentation skills (Kuhn 1991; Means and Voss 1996) or to measure construction of knowledge in a single activity designed to trigger argumentation (Kuhn et al. 1997; Schwarz et al. 2003). We were aware of the growing criticism concerning argumentation skills; for example, Schwarz and Glassner (2003) have pointed to the very high sensitivity of arguments to the context in which they are expressed. Then, some of the criteria proposed to evaluate argumentative aspects of written texts could be the wrong ones. We certainly consider this possibility and definitively did not try to identify the acquisition of skills in the Kishurim program. Number of reasons, and counter-arguments, are for us relevant *indicators* rather than *criteria*. In other words, we knew that an indicator such as the number of reasons invoked might not be adequate as a criterion for evaluating an essay following argumentative activities. Is it justified for detecting traces of dialogical activities in following activities?

Were we actually trying to use inadequate paradigms to detect pre-defined argumentation skills we thought to emerge within an argumentative activity designed as a social learning event? And indeed, the use of Digalo in synchronized argumentative discussion brings to the fore capabilities and know-hows that seem very different from pre-defined argumentative skills (like generating claims and reasons). In Digalo discussions, students engaged in dialogic argumentation and "communicative actions" in practice. These considerations allowed us to consider *new approximations* concerning how we valued Digalo discussions in a way that was *more relevant to our own principles and to the kinds of discussions that developed within Digalo*. We recognized how Goodwin and Heritage (1990) saw the discussion as a context that functioned to "produce a coherent and intelligible course of action" (p. 288). What makes sense in Digalo discussions is how *I* (as a participant in the discussion) relate *my* action to what was said before and might be said later on (Goodwin and Heritage 1990, p. 288). It is in this very context that opinion making and construction of claims, reasons and argument takes place. Their generation expresses dialogism. However, the final essays are not disconnected from this dialogic activity; they are understood by the teacher and the students as a natural continuation of Digalo discussions. And indeed, in the analyses of dialogues we presented, we identified the instillation of the principles on which Kishurim relied and found reasonable interpretations for the improvement of openness, coherence and decisiveness.

Finding traces of dialogic activities in essays is then justified, as well as the comparison of the initial and the final essay. This comparison points to two possible and somehow conflicting directions. The first one concerns a reflective effort to observe activities during the experiment, and to observe some of its products to elaborate novel evaluation tools that are more adequate to reflect changes during the experiment. The second direction is that the program should be improved; that is, we did not fully reach the objectives that we aimed to attain. Such an approach is typical in design-based research.

As for the first direction, we analyzed the initial and final written essays to point at significant differences. The first difference concerned the number of perspectives. The improvement in the number of perspectives is certainly a crucial criterion of the program, since this number conveys an important part of critical dialogue. It reflects the high number of shifts in perspectives exemplified in the dialogue we presented. The same parallelism between dialogue and essays concerned opinion decisiveness. In an example of activity with Digalo that we described, we showed that decisiveness was reflected in activities mediated by the Digalo tool, and that participants proceeded from expression of a personal viewpoint, checking of this viewpoint when confronted with peers' different viewpoints or texts provided by the teacher, and expression of a personal revised viewpoint, often generalized as compared to the initial viewpoint. This result is highly desirable for us as researchers and educators; it shows that the principle of *autonomy* in 'Kishurim' yielded outcomes.

To the same initial attempts to find new methodological tools to adjust measurement to the Kishurim principles, we should add current efforts we pursue but do not report here: collaborative skills such as referring to peers to contribute to the elaboration of a new idea, or knowledge construction during interaction. The higher coherence we found in the post-test is not elaborated here, but we have protocols where students raise arguments with inner contradictions, and are challenged by their peers that lead them to repair these flaws. The criterion of coherence seems then to convey the enactment of both principles of *commitment to critical dialogue* and *collaboration* principles. As shown in the dialogues, *non-intrusive mediation* by the teacher is also necessary.

The absence of effect concerning the number of reasons and claims may point to the fact that students are more inclined to open new perspectives, to be more decided in the

elaboration of their opinions and to be more coherent in their essays, but are less motivated to produce more arguments at any price. The motivation to express their own views *now* seems more important to them than expressing the truth, collectively discussed and reached. And indeed, the discussion we analyzed uncovered highly reasoned discussions because giving reasons helped them in constructing and defending opinions.

Although such an explanation is certainly reasonable, we also considered a second direction: possible shortcomings in the program. Of course the small number of students may reasonably be the very cause for the absence of effect concerning the number of reasons and claims. In fact, we consider that with a larger number of participants effects might be detected—although weaker than for perspectives, decisiveness, and coherence. However, as researchers, we are committed to a critical approach towards our own activities. We already mentioned that among the ten discussions mediated by the Digalo tool, in only two of them (including the one we presented) could students fully express their viewpoints in an autonomous way, compare them with other viewpoints, texts, and peers, and subsequently revise them. In other activities, the argumentative activity was scaffolded by the teacher through questioning techniques, with the students often being in the position of reactants. We subsequently suggested to the teacher to design his activities in a way that was more explicitly directed to the mediation of argumentation. We proposed meta-argumentative techniques ("Do you agree with X?", "You said the contrary 2 min ago?" etc.) rather than argumentative interventions to preserve the autonomy of the discussants. The "failure" of the program according to some criteria induced discussions between the teacher and the research and development team. The same teacher, like many others, continues teaching within our ongoing Kishurim program and continuously reflects with us on past and future practices.

We are currently undertaking a comprehensive analysis of the emergence of new practices and norms during the 24 lessons of the course. In this analysis, we trace how understandings co-constructed in particular activities are actualized (or not) in further ones; we also describe the role of the teacher in these shifts in practices. This analysis is beyond the scope of the present article, but we are already able to suggest that the comparison of the essays gives a partial picture of the evaluation, and that focusing on dialogues only does not give a sufficient perspective about the evaluation of the program. We also describe the important educational role of individual essay writing. In the present paper we showed that the Kishurim program dedicated to dialogism was successful to some extent, and that the tools for its evaluation stemmed from previous studies and evolved dynamically at the same time. This interplay provides opportunities to improve the program and to yield new tools for evaluation that are more adapted to the multifaceted character of dialogism.

Acknowledgments We would like to thank Dan Suthers and the three reviewers for their precious comments that considerably improved the quality of the paper. We would also like to thank the Ziv School and her principal, Maanit Levinger for creating optimal conditions for implementing the Kishurim project. We are deeply indebted to Edith Mani-Ikan for her help in the research and to Yoram Haim, the teacher who designed the History course and implemented it in his class.

References

Andriessen, J., Baker, M., & Suthers, D. (2003). *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments*. The Netherlands: Kluwer.

Bakhtin, M. (1981). *The dialogic imagination: Four essays*. Austin: University of Texas Press (M. Holquist, Ed.; C. Emerson & M. Holquist, Trans.).

Bereiter, C. (2002). *Education and mind in the knowledge age*. Hillsdale, NL: Erlbaum.
Brown, A. L. (1992). Design Experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences, 2*, 141–178.
Buber, M. (1923). *Ich und Du*. Frankfurt: Rütten & Loening.
Cazden, B. C. (1988). *Classroom discourse: The language of teaching and learning*. Portsmouth, NH: Heinemann.
Cobb, P., Stephan, M., McClain, K., & Gravemeijer, K. (2001). Participating in classroom mathematical practices. *The Journal of the Learning Sciences, 10*(1&2), 113–164.
Glassner, A., & Schwarz, B. B. (2005). The role of floor control and of ontology in argumentative activities with discussion-based tools. In T. Koschmann, D. D. Suthers, & T. W. Chan, (Eds.), *Proceedings of CSCL 2005. Computer support for collaborative learning: The Next 10 Years!* (pp. 170–179). Mahwah, NJ: Lawrence Erlbaum.
Goodwin, C., & Heritage, J. (1990). Conversation Analysis. *Annual Review of Anthropology, 19*, 283–307.
Habermas, J. (1972). *Knowledge and human interests*. London: Heinemann Educational Books.
Heidegger, M. (1996). *Being and time: A translation of Sein und Zeit (J. Stambaugh, Trans.)*. Albany, NY: SUNY Press (Original work published in 1927).
Kuhn, D. (1991). *The skills of argument*. Cambridge: Cambridge University Press.
Kuhn, D., Shaw, V., & Felton, M. (1997). Effects of dyadic interaction on argumentative reasoning. *Cognition and Instruction, 15*, 287–315.
Lakkala, M., Lallimo, J., Hakkarainen, K. (2005). Teachers' pedagogical designs for technology-supported collective inquiry: A national case study. *Computers and Education, 45*(3), 337–356.
Means, M. L., & Voss, J. F. (1996). Who reasons well? Two studies of informal reasoning among children of different grade, ability and knowledge levels. *Cognition and Instruction, 14*(2), 139–179.
Scardamalia, M. A., & Bereiter, C. (1994). Computer support for knowledge building communities. *The Journal of the Learning Sciences, 3*, 265–283.
Schwarz, B. B. (2003). Collective reading of multiple texts in argumentative activities. *The International Journal of Educational Research, 39*, 133–151.
Schwarz, B. B. (2005). Do EU funded projects enable collaboration between scientists? The case of R&D in web-based Collaborative Learning Environments. *Computers and Education, 45*, 375–382.
Schwarz, B. B., & Glassner, A. (2003). The blind and the paralytic: fostering argumentation in everyday and scientific issues. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to Learn: Confronting cognitions in computer-supported collaborative learning environments* (pp. 227–260). Dordrecht: Kluwer.
Schwarz, B. B., & Glassner, A. (2007). Designing CSCL argumentative environments for broadening and deepening understanding of the space of debate. In R. Säljö (Ed.), *Information and communication technology and the transformation of learning practices*. Dordrecht: Kluwer (in press).
Schwarz, B. B., Neuman, Y., Gil, J., & Ilya, M. (2003). Construction of collective and individual knowledge in argumentative activity: An empirical study. *The Journal of the Learning Sciences, 12*(2), 221–258.
Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher, 27*(2), 4–13.
Suthers, D. (2003). Representational guidance for collaborative inquiry. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to Learn: Confronting cognitions in computer-supported collaborative learning environments* (pp. 27–46). Dordrecht: Kluwer.
Suthers, D., & Hundhausen, C. (2003). An empirical study of the effects of representational guidance on collaborative learning. *The Journal of the Learning Sciences, 12*(2), 183–219.
van Bruggen, J. M., & Kirshner, J. M. (2003). Designing external representations to support solving wicked problems. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (pp. 177–204). Dordrecht: Kluwer.
van Diggelen, W., Overdijk, M., & De-Groot, R. (2005). *'Say it out loud in writing': A dialectical inquiry into the potentials and pitfalls of computer supported argumentative discussions*. Paper presented at CSCL 2005; May 30 – June 4, 2005, Taipei, Taiwan.
Wegerif, R. (2006). A dialogic understanding of the relationship between CSCL and teaching thinking skills. *The International Journal of Computer-Supported Collaborative Learning, 1*, 143–157.
Wineburg, S. S. (1991). On the reading of historical texts: Notes on the breach between school and academy. *American Educational Research Journal, 28*, 495–519.