Contemp. Educ. Psychol. 33 (2008) 382–402

Contents lists available at ScienceDirect

Contemp. Educ. Psychol.

journal homepage: www.elsevier.com/locate/cedpsych

Effects of argumentation on group micro-creativity: Statistical discourse analyses of algebra students' collaborative problem solving ☆

Ming Ming Chiu [1]

Graduate School of Education, 367 Baldy Hall, State University of New York at Buffalo, Buffalo, NY 14260, USA

ARTICLE INFO

Article history:
Available online 30 June 2008

Keywords:
Cooperative learning
Social interactions
Hierarchical linear modeling
Time series analysis
Quantitative discourse analysis

ABSTRACT

The micro-time context of group processes (such as argumentation) can affect a group's micro-creativity (new ideas). Eighty high school students worked in groups of four on an algebra problem. Groups with higher mathematics grades showed greater micro-creativity, and both were linked to better problem solving outcomes. Dynamic multilevel analyses of the groups' 2951 turns of conversation statistically yielded 53 watersheds (breakpoints) that separated 36 high micro-creativity time periods from 37 low ones. Group member actions within the last two speaker turns also influenced micro-creativity. Compared to agreements, recent disagreements yielded 11% greater micro-creativity. Students who behaved rudely (rude disagreements, commands) showed less micro-creativity (−15% and −9%) than those who behaved politely (polite disagreements, questions/statements). In a special case, after a wrong idea, rude rather than polite disagreements raised groupmates' micro-creativity (+60%). Teachers might encourage students to evaluate ideas carefully, speak politely, and avoid impulsive responses to rude behaviors.

☆ This research was partially funded by a Spencer Foundation research grant, a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CUHK 4117/04H), and a Chinese University of Hong Kong direct grant. The research assistance of Yik Ting CHOI and the helpful comments by Mavis Wu Jing HE and Sung Wook JOH are appreciated.

E-mail address: mingmingchiu@gmail.com

[1] This study began while the author was at the Chinese University of Hong Kong.

0361-476X/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.cedpsych.2008.05.001

1. Introduction

Past research on the development of original ideas that are useful or influential (*creativity*) has largely focused on individuals (Mumford & Gustafson, 1988; Sternberg & Lubart, 1999). However, the explosion of information and specializations will increasingly require teams with diverse skills and knowledge to create innovations (*group creativity*; Dunbar, 1997). For example, groups are more likely than the best individual members to solve some problems (e.g., letters-to-numbers problems, Laughlin, Hatch, Silver, & Boh, 2006). Recently, more researchers have studied group creativity in organizations (West, 2000) and collaborative learning (Johnson, Johnson, & Smith, 2007). However, researchers have not systematically examined how recent group processes (*micro-time context*) affect the likelihood of creative moments (*group micro-creativity*).

In this study, a step is taken in this direction by analyzing how group processes such as argumentation affect micro-creativity during 20 groups of students' algebra problem solving. In this paper, creativity refers to the "small c" creativity of ordinary people in daily life, not the "big C" creativity of new knowledge or products that substantially affect society (Gruber & Wallace, 1999). Hence, micro-creativity is defined as an expressed idea that is new relative to the group members' experiences.

This study contributes to the research literature in four ways. First, hypotheses are introduced regarding how argumentation and other group processes help create a micro-time context that might influence micro-creativity. Second, statistically-identified watershed breakpoints divide a group's problem solving actions into distinct time periods of high micro-creativity and low micro-creativity. Third, hypotheses regarding how argumentation processes might affect micro-creativity are tested with a new statistical discourse analysis method (Chiu & Khoo, 2005; cf. Fleming, Shire, Jones, Pill, & McNamee, 2004; Pirie, 1991). Fourth, the micro-creativity hypotheses are tested for significant differences across groups or across time periods. By understanding the group processes that affect micro-creativity, group members can work together more creatively.

2. Group processes and micro-creativity

Past research shows that some group characteristics and processes (diversity, argumentation) aid collaborative problem solving while others (public self-image [face], status concerns) hinder it. In this study, several hypotheses are introduced regarding how recent group processes might influence micro-creativity, which in turn might affect problem solving success. Specifically, diversity and argumentation might increase micro-creativity, while face and status concerns might reduce micro-creativity.

2.1. Group processes that aid micro-creativity

2.1.1. Diversity

Groups with diverse views can create more ideas, representations, justifications, and solution proposals, especially when group members value one another's diverse contributions (Larson, 2007; Paulus & Brown, 2003; Stasson & Bradshaw, 1995; Swann, Kwan, Polzer, & Milton, 2003). Specifically, group diversity in nationality or ethnicity often increases the number of perspectives, number of ideas, and quality of ideas (see Fig. 1; McLeod, Lobel, & Cox, 1996). Group members' diverse views can also help them justify the validity of an idea, identify flaws, and revise them to create new ideas (De Lisi & Goldbeck, 1999; Orlitzky & Hirokawa, 2001). Hence, group diversity can enhance micro-creativity.

Group members with diverse views can also build on one another's ideas through processes such as sparked ideas, jigsaw pieces, and creative misinterpretations (Paulus & Brown, 2003). Comments by one person (e.g., a key word) might spark another person to activate related concepts in his or her semantic network and propose a new idea (Nijstad, Diehl, & Stroebe, 2003). Like fitting jigsaw pieces together, group members also can put together ideas to create a new idea (Milliken, Bartel, & Kurtzberg, 2003). Lastly, a person might also misinterpret another person's idea to create a new idea (Chiu, 1997). Thus, even wrong ideas can be kindling for new ideas. This contrasts with the view that people only build on correct ideas and that wrong ideas lead the group astray. Hence, both correct and incorrect, new ideas might aid micro-creativity.

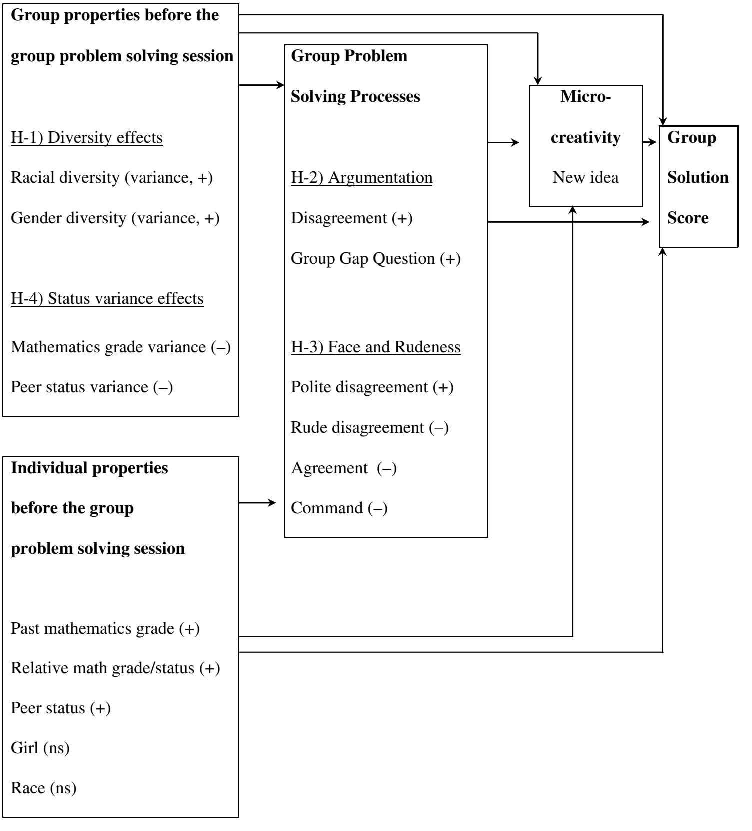

Fig. 1. Model of the effects of student and group properties before and during the group problem solving process on the outcome variables, correct contributions and group solution score (symbols in parentheses indicate expected direction of relationship with the outcome variables: positive [+], negative [−], or not significant [ns]).

Successful collaborative problem solving often involves argumentation in the cognitive/problem content space (Teasley & Roschelle, 1993), a social process by which people explain and justify their own views to convince both themselves and others (Cobb, 1995; Jehn & Mannix, 2001). During this process, group members evaluate one another's ideas, recognize problems, and justify their views (Cobb, 1995; Kuhn, Shaw, & Felton, 1997).

2.1.2. Evaluations

Evaluations characterize how a person assesses the previous speaker's action and problem solving approach (*functional theory of group decision-making*, Janis, 1989; Orlitzky & Hirokawa, 2001). For example, Bob says "two plus three is six." Ann can agree ("right"), use a neutral action ("louder, can't hear you"), disagree ("no, you're wrong"), or change the topic ("I'm hungry"). While agreements encourage continuation of the current problem-solving trajectory, disagreements and changes of topic (ignoring the previous action) try to change it (West, 2000).

Group members can recognize problems or difficulties (perturbations), express them through disagreements or questions, and try to address them (*socio-cognitive conflict* theory, Buchs, Butera, Mugny, & Darnon, 2004; Doise, Mugny, & Perret-Clermont, 1975; Piaget, 1985). Piaget (1985) defined two types of perturbations: (a) obstacles, often expressed through negative feedback (disagreement) and (b) lacunae, gaps in understanding, often expressed through questions.

2.1.3. Disagreements

Disagreements might increase micro-creativity both directly and indirectly. A group's collective attention and diverse perspectives can help it identify more flaws compared to individuals (Milliken et al., 2003). By identifying obstacles to be overcome (e.g., "no, that's wrong, two plus three isn't six"), a disagreement might directly stimulate micro-creativity, especially from social loafing students who no longer rely on other group members to solve the problem. Furthermore, a disagreement (even if wrong) often stimulates the attention of group members (De Dreu & West, 2001; Nemeth & Rogers, 1996). Thus, a disagreement might help them consider more aspects of problem from more perspectives to increase micro-creativity indirectly.

Furthermore, disagreements might generally encourage group members with minority views to express their ideas, especially after agreements and repetitions of an existing idea suggest a majority view (Nemeth & Chiles, 1988). Thus, a disagreement by another member, regardless of its validity, legitimizes different opinions, freeing all group members to express their ideas, including those unrelated to the specific disagreement (Nemeth & Chiles, 1988). This greater legitimacy of different opinions can aid micro-creativity.

2.1.4. Questions

Meanwhile, a question might indicate an individual gap or a group gap, and hence yield different effects on micro-creativity [to aid reader comprehension, "question," "command," and "statement" are used rather than "elicitation," "directive," and "informative" (Tsui, 1992)]. A person asking a question shows a gap in his or her understanding (Tsui, 1992). For example, Ann asks, "how do we find the speed?" If another group member can fill this gap with an old, previously expressed idea (e.g., Bob says "Six miles divided by two hours is three miles per hour"), the question is an individual gap question. Thus individual gap questions encourage review of old ideas instead of creation of new ideas. In contrast, no group member knows the answer to a group gap question, which motivates the need for micro-creativity and points to a direction for creating a new idea (see Fig. 1, middle column). Thus, individual and group gap questions might show opposite effects on micro-creativity.

In short, two hypotheses regarding diversity and argumentation are proposed.

H-1. Groups with more diversity (race, gender) show higher micro-creativity.

H-2. Argumentation (disagreement, group gap question) aids micro-creativity.

2.2. Collaborative problem solving actions that hinder micro-creativity

Although disagreements might aid group problem solving (socio-cognitive conflict theory), rude and polite disagreements yield different effects according to *politeness* theory (Brown & Levinson, 1987). Polite disagreements likely facilitate collaborative problem solving, but rude disagreements can hinder it, especially during status struggles. When arguments spill over from the problem content space into the social relational space (Barron, 2003; Jehn & Mannix, 2001), groups members might protect their *face* (defined as a person's public self-image) rather than further the problem solving (Brown & Levinson, 1987). Furthermore, these face concerns might be aggravated by status differences.

2.2.1. Face and rudeness

Each type of evaluation has implications for both the problem solving (as noted above) and for the previous speaker's face (Chiu, 2000a). Evaluations range from polite to rude: agreement, neutral, change of topic, and disagreements (*politeness* theory, Brown & Levinson, 1987). Consider Bob's utterance again, "two plus three is six." If Ann agrees with Bob ("right"), she supports him, promotes his face, and enhances their social relationship (Brown & Levinson, 1987). Thus, members often create common ground and solidarity by repeating shared information (Stasser & Birchmeier, 2003) and by spontaneously reciprocating positive affective displays (e.g., eye contact) to indicate agreement

(Burgoon, Dillman, & Stern, 1993). Greater shared understanding and social solidarity is often linked to greater group efficiency and effectiveness (e.g., Civettini, 2007).

However, desire for agreement might reduce expression of new ideas. A group member with both supportive information and conflicting information tends to agree and provide supportive information, which is likely more salient and hence more likely to be activated quickly (Larson, Christiansen, Abbott, & Franz, 1996). Even if both pieces of information are activated, the common information tends to be more credible and compelling because socially validated consensus is often correct (Chaiken & Stangor, 1987). Based on probabilities, a common idea shared by many people is more likely to be voiced and supported than an unusual idea held by a few people (Stasser & Birchmeier, 2003). Hence, agreements support face but reduce expression of new ideas.

In contrast, other actions do not support face. Neutral actions include many discourse management or meta-discourse actions (e.g., "louder, can't hear you"). Although changes of topic ("I'm hungry") can be neutral, they are rude if the previous speaker (Bob) expects a response (e.g., Bob asks the question, "is two plus three, six?"). If Ann says "I'm hungry" after Bob's question, she either ignores him or does not listen to him, both of which are rude. Lastly, disagreements (e.g., "you're wrong") can threaten face by lowering public perception of the previous speaker's (Bob's) competence (Brown & Levinson, 1987).

When a person disagrees (e.g., Ann), the target person (Bob) ideally tries to understand and use the criticism to either justify his original idea or create a new idea. However, people are prone to defending their ideas (Paulus, 2000). Hence, the threat to Bob's face may encourage him to react impulsively and retaliate emotionally (*face attack*, "no, I'm not! You're wrong. You're always making mistakes . . .."; Tracy & Tracy, 1998). Thus, rude disagreements threaten face, reduce psychological safety, and escalate interpersonal conflict (Turner & Horvitz, 2001). In the worst case, a spiral of rude disagreements can kill the collaboration. Even if the collaboration survives after a rude disagreement (or some other rude action, e.g., insult), group members might withhold their ideas rather than risk losing face (Mulryan, 1992). In contrast to socio-cognitive theory, rude disagreements might reduce micro-creativity (see Fig. 1, middle column).

To avoid threatening Bob's face, Ann might go to the opposite extreme, withhold her disagreement, and publicly agree. By doing so, Ann enhances her social relationship with Bob at the expense of their problem solving. Such false agreements allow errors to persist and potential new ideas to remain unspoken (see Fig. 1, middle column). For example, teenage girls often avoid disagreeing with one another (Tudge, 1989), and tutors often do not point out their students' errors (Person, Kreuz, Zwaan, & Graesser, 1995). Thus, false agreements, as well as true agreements, might hinder micro-creativity.

Avoiding the extremes of rude disagreement and false agreement, Ann can disagree politely (with redress) to reduce the threat to Bob's face while maintaining problem solving integrity (Brown & Levinson, 1987). Instead of "no, you're wrong," Ann can disagree politely, "If we add two and three, we don't get six." The polite disagreement both reduces blame and creates common ground. First, Ann uses the hypothetical "if," thereby distancing the idea from reality. Second, she does not refer to Bob, (no "you") thereby avoiding assignment of blame. Third, Ann uses the passive voice, "is multiplied," rather than the active voice, to hide causal agency and responsibility. Lastly, she uses the passive circumstantial verb "get," thereby implicating agency in external conditions.

Ann's polite disagreement creates common ground by repetition and shared positioning. By repeating Bob's computation, "two . . . three . . . six," Ann suggests that she shares his understanding of the arithmetic expression. Also, Ann uses shared positioning, specifically the first person plural pronoun "we," to claim common cause with Bob.

Ann's polite disagreement supports her relationship with Bob, so he is less likely to retaliate. Instead, Bob is more likely to try to understand Ann's criticism, recognize the flaw, and correct it to create a new idea (Chiu & Khoo, 2003). Indeed, the benefits of polite disagreements are so strong that it is the accepted norm among peers, as lack of redress during a disagreement is noticeably rude and unacceptable (Holtgraves, 1997). In short, polite disagreements might support both social relationships and micro-creativity (see Fig. 1, middle column).

Like disagreements, commands are often rude. Unlike questions, commands dictate a specific course of action without any need for further explanation or justification (Tsui, 1992). As commands demand action from the target listener, they impinge on the target listener's freedom and are less polite than questions or statements (Brown & Levinson, 1987). Furthermore, the target listener(s) is expected to comply rather than create new ideas. Thus, commands might hinder micro-creativity (see Fig. 1, middle column).

2.2.2. Status

Status differences can hinder collaborative problem solving through individual pursuit of high status via status struggles (*status characteristics* theory, Bales, 2001; Gersick, 1988) or through the greater influence of high status members (Dembo & McAuliffe, 1987). Cohen (1994) defined status as "an agreed-on rank order where it is generally felt to be better to be high than low rank" (p. 23).

As a higher status person often receives more group resources and attention, people often compete for higher status (status struggles), especially if no status hierarchy has been established (Bales, 2001; Gersick, 1988). During status struggles, intentional rude actions (*face attacks*; e.g., "three plus two is *obviously* five, not six") can enhance one's own face at the expense of a competitor's face (Tracy & Tracy, 1998). As noted earlier, rude disagreements can hinder micro-creativity.

After a status hierarchy has been established, status affects the expectations of individual group members (Berger, Cohen, & Zelditch, 1972; Cohen, 1994). Status is linked to the expectation of competencies for the current activity (*expectation states* theory, Berger et al., 1972). High status is conferred on group members who are expected to contribute positively to a desired outcome. These expectations create different opportunities to perform and receive rewards. Members can selectively invite and defer to high status members' opinions while discouraging, undervaluing, or outright ignoring lower status members' ideas. By doing so, members enact their expectations of high status members dominating the interaction to the detriment of lower status members during the initial stages of group activity (Hackman & Johnson, 2000; Larson et al., 1996; Turner, 2005) (Mathematics anxiety has a parallel effect, as group members with high mathematics anxiety are less likely to contribute successfully during the initial stages of the group activity (Ashcraft & Krause, 2007; Ma & Xu, 2004)).

Dominated by higher status group members, these early group interactions create initial group member identities and group discourse norms. Afterwards, group members often prefer and seek self-verifying evaluations of their initial identities from others (Swann, 2005; Swann et al., 2003). Furthermore, the patterns of early discourse often become the preferred discourse norms (Sfard, 2007; Yackel & Cobb, 1996). As group members value and prefer supporting previously discussed or shared information rather than introducing unshared information, high status members' influence further increases in over time (Stasser & Taylor, 1991). Hence, lower status group members might enact these lower expectations and identities, withhold their ideas, and reduce the group's overall micro-creativity.

Thus, greater status differences might increase the incentives for status struggles and yield greater status effects, both of which might reduce micro-creativity (see Fig. 1, left column). For collaborative problem solving among students, the primary status characteristic is often past achievement (which also affects mathematics anxiety, Ashcraft & Krause, 2007; Ma & Xu, 2004), but group members might also use diffuse status characteristics (e.g., gender, race) to make assumptions about a person's competence (Cohen, 1982; Webb, 1984).

H-3. Polite disagreements aid micro-creativity, but rude disagreements, false agreements, and commands hinder it.
H-4. Greater status differences reduce micro-creativity.

2.2.3. Successful vs. unsuccessful groups

The above four hypotheses suggest that recent group characteristics and actions (micro-time context) can influence micro-creativity. As micro-creativity might be linked to problem solving success, successful groups might tend to show more characteristics and recent actions that aid micro-creativity and fewer ones that hinder micro-creativity. Whether these group characteristics or actions show different effects on micro-creativity across groups or across time periods within a group's problem solving session remain open questions.

In earlier studies of mathematics discourse (e.g., Chiu & Khoo, 2003), the variation in student actions during group problem solving occurred mostly at the speaker turn level rather than at higher levels (e.g., group or classroom). Thus, this study focuses on simpler, proximal analyses of speaker turns, time periods, and groups (leaving more complex, distal analyses involving classroom and school differences for future studies, e.g., openness and supportiveness of the classroom environment, classroom argumentation norms, and so on; Sfard, 2007; Yackel & Cobb, 1996). As past studies showed that gender, past achievement, and social status (reflecting social skills) can affect mathematics performance (e.g., Ding, Kim, & Richardson, 2007; Wentzel & Watkins, 2002), they were added as control variables.

3. Methods

To test the model in Fig. 1, student questionnaires were collected, and groups of students were videotaped while solving an algebra problem. Then, the data were analyzed at the group, time period, and speaker turn levels to model problem solving outcomes and processes (The same transcripts were used in the Chiu & Khoo (2003) study, but it addressed a different research question: how do group processes affect student evaluations of one another's ideas? Specifically, that study used a different outcome variable "agree with previous speaker" [vs. not agree]. In contrast, the outcome variable in this study is micro-creativity [new idea]).

3.1. Participants

The participants attended four ninth grade algebra classes in an urban US high school, which had an overall school score at the 40th percentile (maximum = 100; California Department of Education, 2005). Eighty-seven students were asked to answer a status survey and to be videotaped. Of the 87 students, 7 (or 8%) declined to participate (of these 7 students, 4 were girls and 3 were boys. The mean of these non-participating students' mathematics grades in the previous semester was 77% compared to a mean of 82% for the participating students). There were 40 girls and 40 boys participating. Their races were 12 Asian, 27 Black, 28 Hispanic, and 13 White.

These students worked in groups of four, had not previously worked together, and had not received any group work training. They attended the same algebra course for seven months and were likely aware of one another's mathematics abilities (as classmates chat and see one another's classroom behaviors). Thus, group members' relative mathematics abilities were more likely to have a primary status effect. Likewise, diffuse status characteristics such as gender and race were likely to have smaller effects on their interaction compared to that of strangers (Cohen, 1982; Webb, 1984).

3.2. Procedure

All 80 algebra students who agreed to participate answered four questions regarding status: "Who are 3 classmates you would most like to hang out with?" "Who are 3 classmates you would choose for your group to learn the most math?" "Name 3 classmates who are the easiest for you to talk with outside of school work." "Name 3 classmates that could help you the most with a super hard math problem."

Later, their teacher presented the following problem in their algebra classes:

> "You won a cruise from New York to London, but you arrive 5 hours late. So, the ship left without you. To catch the ship, you rent a helicopter. The ship travels at 22 miles an hour. The helicopter moves at 90 miles an hour. How long will it take you to catch the ship?"

As advocated by cooperative learning researchers (e.g., Cohen, 1994; Johnson & Johnson, 1994), this problem was challenging for these students and had multiple solution methods. The classes had studied equations with single variables, and the teacher used the above problem to introduce them to a new unit on algebraic equations with multiple variables. Hence, the students had not yet learned any procedures for solving this problem during class. Furthermore, the problem involved complicated mathematics relationships, non-trivial combinations of multiple operations, and a non-integer solution. One solution is to equate the cruise ship's and helicopter's distance computations (22 mph [Time + 5 hours] = 90 mph $\times$ Time), to obtain 1.618 hours or 1 hours 37 minutes.

The students worked in groups of four for 30 min. They had pens, paper, and calculators. There were six to seven video cameras in a classroom, one following the teacher and one for each group of students.

Likewise, the teacher and each group of students had their own microphone and audiotape recorder to backup the video recordings. The videotape data were transcribed, coded, and analyzed.

3.3. Variables

See Table 1 for summary statistics and descriptions of variables. Using a similar set of data from a pilot study, two research assistants were trained to transcribe and code the videotapes. Each transcript was divided into sequences of words or actions by a group member bracketed by the words of other group member(s) (*speaker turns*). Turns unaccompanied by words (e.g., writing "3×40") were also counted as speaker turns. Blind to the study's hypotheses, the research assistants coded each speaker turn from the videotape on to a transcript, maintaining a log of each videotape to facilitate their coding. To compute the inter-rater reliability, Krippendorff's α was used (2004). Unlike other inter-rater reliability measures, Krippendorff's α applies to any number of coders, any number of categories or

Table 1
Summary table of group level variables

Group variable	Mean	*SD*	Description
Outcome variables			
Solution score	1.90	1.25	Score of a group's final solution (0–3)
Micro-creativity	0.40	0.17	An idea that has not been mentioned earlier during the group problem solving, during class, or in the textbook
Before problem solving			
Classroom_1	0.25	0.44	Binary variable for students in classroom 1. Baseline classroom is classroom 4
Classroom_2	0.30	0.46	Binary variable for students in classroom 2
Classroom_3	0.25	0.44	Binary variable for students in classroom 3
Girl	2.00	0.65	Number of girls in each group (0 indicates all boys)
Asian	0.60	0.50	Number of Asians in each group
Latino	1.40	0.75	Number of Latinos in each group
White	0.65	0.49	Number of Whites in each group
Mathematics grade	82	7	Mean of all students' last semester's mathematics grades within a group
Social status	23	8	Mean number of times a student's name appeared in classmates' answers to the following questions. Who are 3 classmates you would most like to hang out with? Name 3 classmates who are the easiest for you to talk with outside of school work. This measure is the mean for the group and serves as a proxy for the group's social skills
Measures of status effects			
Math grade variance	101	70	Variance of students' mathematics grade within a group
Social status variance	37	29	Variance of peer status within a group
During problem solving			
Total no. of words	1363	1174	Total number of words during a group's problem solving sessions. The large standard deviation is partially due to five groups that spoke over 2400 words each
Total on-task words	1338	1277	Total number of words spoken on on-task turns during a group's problem solving sessions. See total words note
Number of new ideas	43	33	Number of new ideas that have not been mentioned earlier during the group problem solving session
Wrong idea	0.17	0.10	A flawed idea
Rude disagreement	0.09	0.09	Disagree with the previous speaker without redress
Agreement	0.58	0.10	Agree with the previous speaker
Command	0.06	0.07	A directive that invites a non-verbal response (Sinclair & Coulthard, 1992)

scale values, any level of measurement, any sample size, and incomplete data. Its values range from −1 (maximum disagreement) to 1 (perfect agreement). A value near 0 indicates chance agreement among coders, and a value of 0.7 or higher indicates satisfactory agreement.

The research assistants settled coding disagreements by consensus, if possible. They could not agree in 19 cases, so the author made the final coding decision. Due to poor sound quality, 49 speaker turns could not be coded. These turns were coded as *missing* and inspected with adjacent outcome variables and explanatory variables for significant correlations. As they did not correlate significantly with other variables, omitting them was not likely to affect the results.

3.3.1. Speaker turn variables

In contrast to flat classification schemes that only allow one or two codes for each speaker utterance (e.g., Bales, 2001; Barnes & Todd, 1995; Sinclair & Coulthard, 1992), the research assistants coded each turn along the five dimensions suggested by these turn properties: evaluation of the previous action (agree, politely disagree, rudely disagree, ignore, neutral), knowledge content (old idea, new idea [micro-creativity], no problem content [null, e. g., "hungry?"]), validity (right, wrong, null), justification, and invitational form (question, command, statement; Chiu, 2000b). An idea was coded as new if it was not in the problem statement, not in the textbook, and not mentioned earlier by group members, classmates or the teacher during class. Note that some of these codes require assessing interactions or relationships across speaker turns (*relational* measures, e.g., a speaker's evaluation of the previous speaker's idea). See Tables 1–3, and Chiu (2001) for summary tables, the coding decision trees, and further details.

3.4. Data analysis

The link between group micro-creativity and problem outcome was analyzed at the group level. Then, crucial watersheds that separate high micro-creativity time periods from low micro-creativity time periods were statistically identified. After incorporating these time periods into the analytical model, micro-creativity in each speaker turn was analyzed.

3.4.1. Predicting solution score at the group level

Solution score is an ordered variable (with possible values of 0, 1, 2, or 3; not a continuous one). As ordinary least squares (OLS) regressions would bias the standard errors, ordered Logit/Probit was used (Kennedy, 2004).

The following independent variables were entered. First, classroom identification binary variables were entered to control for classroom effects. Time constrains the direction of causality, so group processes cannot affect a priori group characteristics. Hence, characteristics of group members were entered into the regression before entering group process properties. The variables "total number of words" and "total number of on-task words" were added to control for the amount of talk by each group. The order of entry was: gender, race, mean mathematics grade, mathematics status, social status, gender variance, race variance, mathematics grade variance, mathematics status variance, social status variance, words, on-task words, and number of new ideas.

Table 2
Coding of a classroom discourse segment along five dimensions: (1) evaluation of the previous action (EPA: agreement [+], polite disagreement [−], rude disagreement [---], ignore/new topic[*]), (2) knowledge content (KC: new idea [I], repetition [R], null academic content [N]), (3) validity (right [√], wrong [X], null academic content [N]), (4) justification (justification [J], no justification [], null academic content [N]), and (5) form of invitation to participate (IF: (command [!], question [?], statement [_.])

Person	Action	EPA	KC	Validity	Justify	IF
Ann	We need to do three plus two	*	I	√	[]	!
Kyle	Three plus two is, um,	+	R	√	[]	_.
Jean	Six	+	I	X	[]	_.
Bob	No, you're wrong	---	N	√	[]	_.
Kyle	If we do three plus two, don't we get five because it's like three and one is four and one is five?	–	I	√	J	?
Ann	Yep	+	N	N	N	_.

Table 3
Summary table of speaker turn variables

Speaker turn level variable	Overall		Successful		Unsuccessful		Min.	Max.
	Mean	*SD*	Mean	*SD*	Mean	*SD*		
Unsolved[a]	0.56	0.50	0	0	1	0	0	1
Micro-creativity	0.29	0.46	0.36	0.48	0.24	0.43	0	1
Before problem solving								
Girl	0.46	0.50	0.42	0.49	0.50	0.50	0	1
Asian	0.14	0.35	0.13	0.34	0.15	0.36	0	1
Latino	0.25	0.43	0.30	0.46	0.21	0.41	0	1
White	0.26	0.44	0.20	0.40	0.30	0.46	0	1
Mathematics grade	86	10	91	8	82	10	64	100
Social status	23	8	28	8	19	5	9	37
Variance of math grade	97	75	65	31	121	89	12	299
Variance of peer status	31	25	53	14	14	19	0	107
During problem solving								
Wrong idea	0.14	0.35	0.15	0.36	0.13	0.34	0	1
Polite disagreement	0.18	0.39	0.19	0.39	0.18	0.39	0	1
Rude disagreement	0.10	0.30	0.06	0.24	0.12	0.33	0	1
Agreement	0.56	0.50	0.63	0.48	0.51	0.50	0	1
Command	0.07	0.26	0.10	0.30	0.05	0.22	0	1

[a] Separate analyses for groups with each solution score showed substantial differences between groups that did and did not solve the problem correctly, and similar results across the latter unsuccessful groups. Thus, *unsolved* was coded as a binary variable (0 or 1) in the turn-level analysis to facilitate interpretation of the results.

A nested hypothesis test (χ^2 log likelihood) checked whether each set of added variables was significant (Kennedy, 2004). Only significant variables were retained. The small sample size ($N = 20$ groups) limited the statistical power of this analysis to identify non-significant results at the group level (power = 0.25 for an effect size of 0.3).

3.4.2. Solving analytical difficulties with dynamic multilevel analysis

Statistical analyses of group processes at the speaker turn level must overcome three difficulties. First, group members' behaviors and effects differ across groups and across time, yielding nested data. Second, outcome variables are often discrete, not continuous. Third, events are often similar to recent events, so variables tend to depend on values from recent turns (serial correlation).

OLS regressions do not address these difficulties. First, OLS often underestimates the standard errors of regression coefficients when applied to nested data (Goldstein, 1995). Second, OLS is statistically inefficient for discrete variables and yields biased results (Kennedy, 2004). Lastly, if the time-series relationships are not modeled properly, the model residuals can be serially correlated, resulting in inefficient parameter estimates and biased estimates of the parameters' standard errors (Kennedy, 2004).

These difficulties were addressed with a statistical discourse analysis tool, dynamic multilevel analysis (DMA, Chiu & Khoo, 2005). DMA was used to identify watersheds (breakpoints), test for group and time period differences, model micro-creativity, test for serial correlation, and model direct and indirect effects.

3.4.3. Breakpoints separate high and low micro-creativity time periods

DMA was used to identify watershed breakpoints that separate high micro-creativity time periods from low micro-creativity time periods (for details, see Chiu & Khoo, 2005). *T*-tests were used to determine whether breakpoint characteristics differed among successful and unsuccessful groups. The small sample size limited the statistical power of this analysis to identify non-significant results (number of breakpoints = 53; power ≈0.60 for an effect size of 0.3).

3.4.4. Effects on micro-creativity

Next, links between explanatory variables and micro-creativity were modeled with DMA via multilevel Logit. As the data had three levels (group, time period, and speaker turn), a multilevel analysis might be needed (specifically a 3-level model). Also, the binary outcome variable (micro-creativity) required a Logit or Probit model. A multi-level Logit variance components model (Goldstein, 1995; Bryk & Raudenbush, 1992) was used to test if the outcome variable, micro-creativity, significantly varied across groups or across time periods with the software MLn (Rasbash & Woodhouse, 1995).

First, a vector of s classroom identification variables was added as control variables: classroom_1, classroom_2, classroom_3 (**S**). The parameter s refers to the number of variables in the vector **S**, in this case, three.

$$\pi_{ijk} = p(y_{ijk} = 1 \mid \mathbf{S}_{00k}, \beta_{00s}) = F(\beta_{000} + \beta_{00s}\mathbf{S}_{00k} + f_{0jk} + g_{00k}) \tag{1}$$

f_{0jk} and g_{00k} represent the deviations of time period j and of group k from the overall mean β_{000}. The probability (π_{ijk}) that event y_{ijk} (new idea) occurs at turn i of time period j in group k is determined by the expected value of y_{ijk} and the Logit link function (F, see Kennedy, 2004). Non-significant variables were removed at each specification.

Then, a vector of t variables was entered at the group level: incorrect group solution, mean number of girls in the group, mean numbers of Asians, Blacks, and Latinos, mean of group members' mathematics grades, mean of group members' mathematics statuses, mean of group members' social statuses, gender variance, racial variance (white vs. non-white group members), variance of mathematics grades, variance of mathematical statuses, and variance of social statuses (**T**). This specification tested the diversity and status differences hypotheses, H-1 and H-4. As with **S**, a Wald test was done on **T**. Then, interaction effects among pairs of significant variables in **T** were tested.

$$\pi_{ijk} = F(\beta_{000} + \beta_{00s}\mathbf{S}_{00k} + \beta_{00t}\mathbf{T}_{00k} + \beta_{ujk}\mathbf{U}_{ijk} + \beta_{vjk}\mathbf{V}_{(i-1)jk} + \phi_{vjk}\mathbf{V}_{(i-2)jk} + f_{0jk} + g_{00k}) \tag{2}$$

Next, u current speaker variables were added at the speaker turn level: gender, race, mathematics grade, mathematics status, social status, correct evaluation, agree, politely disagree, rudely disagree, justify, question, and command (**U**). Likewise, the procedure for **T** was applied to **U**. Each u speaker turn level regression coefficient ($\beta_{ujk} = \beta_{u00} + f_{ujk} + g_{u0k}$) was tested for different effects across time-periods ($f_{ujk} \neq 0$?) and across groups ($g_{u0k} \neq 0$?; Goldstein, 1995). If the regression coefficients differed across time periods or across groups, then a variable at the time period level or group level, respectively, moderated the effect of the explanatory variable.

Using a vector autoregression (VAR, Kennedy, 2004), lag variables were entered for the previous speakers, first lag 1 (indicating the previous turn and denoted −1), then at lag 2 (denoted −2), and so on until none of the variables in the last lag were significant (lag 2 in this case). First, v previous speaker variables were added at the speaker turn level: gender (−1), race (−1), mathematics grade (−1), mathematics status (−1), social status (−1), agree (−1), rudely disagree (−1), politely disagree (−1), correct (−1), wrong (−1), micro-creativity (−1), correct, old idea (−1), wrong, old idea (−1), justify (−1), question (−1), and command (−1) (**V**). As shown in Fig. 1, these variables test the argumentation and rudeness hypotheses (H-2, H-3). The procedure for **U** was applied to **V**. Then, the procedure was applied for lag −2 of the variables in **V**.

An alpha level of .05 was used for all statistical tests. The false discovery rate (FDR) was controlled with Benjamini, Krieger, and Yekutieli (2006) two-stage linear step-up procedure, as computer simulations showed that their procedure addresses this issue better than 13 other methods.

Ljung and Box (1979) Q-statistics were used to test for serial correlation (up to order 4) in the residuals for all 20 groups. If the residuals are serially correlated, the serial correlation can be modeled (see Goldstein, Healy, & Rasbash, 1994, for details).

Based on the multilevel analysis results, the path analysis estimated the direct and indirect effects of the significant explanatory variables separately to compute their total effects (Kennedy, 2004). As time constrains the direction of causality, the explanatory variables were entered in temporal order into the path analysis.

To facilitate interpretation of these results, the effect of each explanatory variable was converted to an odds ratio, and the total effect (direct plus indirect) was reported as the percentage increase or decrease (+X% or −X%) in micro-creativity (Kennedy, 2004). As the underlying distribution was unknown, the above analyses were repeated with multi-level Probit to ensure that the results were not dependent on the Logit distribution. The predictive accuracy of the final micro-creativity model was estimated by comparing the final model's prediction of whether a new idea occurred at each speaker turn in each group ($y_{ijk}{}^{*}$) with the new idea's actual presence or absence (y_{ijk}).

A multilevel analysis has multiple units of analysis, so the statistical power for each one (group, time period, speaker turn) must be computed separately. At the speaker turn level, the sample size is 2951, so the statistical power is over 0.99 even for a small effect size of 0.1. Thus, using many speaker turn-level variables is acceptable. However, higher level units (time period, group, classroom, school, country) are not necessarily representative, so these higher level results require cautious interpretations as they might differ in other higher level unit contexts. As students can change behaviors during another student's speaker turn, modeling students as a level of analysis requires multivariate outcome, multilevel, cross-classification logit/probit, but no implementation of such a method has been shown at this time. A non-dynamic analysis of micro-creativity has been included in Appendix A to highlight its differences with DMA.

4. Results

As expected, groups with higher mean mathematics grades showed higher micro-creativity ($b = .011$, $SE = .001$, $R^2 = .17$), and both were linked to higher solution scores (mean mathematics grade: $b = .23$, $SE = .09$; micro-creativity: $b = 7.37$, $SE = 3.02$; $R^2 = .31$; likelihood ratio $\chi^2 = 15.1$ [df = 2; $p < .01$]; ordered Probit showed similar results). Next, 53 watershed breakpoints divided the groups' problem solving sessions into 36 high micro-creativity time periods and 37 low micro-creativity time periods. Modeling these time periods with DMA, both disagreements and politeness were linked to greater micro-creativity. Krippendorf's α for evaluation of previous actions, knowledge content, validity, justification, and invitational form were 0.93; 0.98; 0.99; 0.91; and 0.91 (corresponding percentages of agreement were 96%, 98%, 99%, 96%, and 96%). Due to space considerations, only the main results are included. All results are available upon request.

4.1. Breakpoints separate high and low micro-creativity time periods

Examination of the 53 breakpoints identified via DMA showed that the ten successful groups had more break points (31), and hence more time periods (41 total; $M = 4.1$, $SD = 1.91$) than the ten unsuccessful groups (22 breakpoints; 32 time periods: $M = 3.2$, $SD = 1.55$). However, the difference was not significant (t-test = 0.263; $p > .05$). In both successful and unsuccessful groups, the number of breakpoints ranged from zero to four, and the number of time periods ranged from one to five. Successful and unsuccessful groups showed no significant differences across any of the above characteristics, across total problem solving time, or the time between breakpoints.

An exploratory analysis showed three types of breakpoints: off-task ↔ on-task transitions, creativity dampeners, and creativity ignitors. Groups transitioned from off-task to on-task or vice-versa at 27 of the breakpoints (15 in successful groups, 12 in unsuccessful groups). At 9 of the breakpoints (5, 4), group member actions sharply decreased micro-creativity (*creativity dampeners*). Creativity dampeners included inadequate explanations, insults ("It's a hundred and ten, don't you know how to times?"), and emotional arguments ("Yes, it is," "No, it's not"). Consider the creativity dampener in the following transcript segment (All names are pseudonyms).

27 Jan: Ship goes five times twenty-two.
28 Ron: [presses calculator keys 5 × 22=] one ten.
29 Jan: A hundred and ten miles. [writes "110"].
30 Ron: The helicopter is ninety miles an hour, so [presses calculator keys 5 × 90=] four fifty.
31 Jan: Four fifty?
32 Xia: Uh-huh.
33 Jan: That don't make sense.
34 Ron: Yes, it does.
35 Jan: No, it doesn't.

36 Ron: [presses calculator keys 5 × 90=] four-fifty
37 Xia: Uh-huh.
38 Jan: But what's four fifty?
39 Ron: That's what you get when, when you do five times ninety, right?
40 Xia: Uh-huh.
41 Jan: But why five times ninety?
42 Ron: Five hours times ninety miles per hour.
43 (9 seconds of silence).

Line 31 is a breakpoint that separates a high micro-creativity time period (lines 27–30) from a low one (lines 31–43). In line 30, Ron's idea to multiply the ship's travel time of 5 hours and the helicopter's speed is inconsistent with the problem situation ("helicopter is ninety miles an hour, so [presses calculator keys 5 × 90=] four fifty"). Afterwards, Jan questions Ron's computation ("four fifty? ... that don't make sense. ... why five times ninety?). However, Ron only repeats the computation procedure without explaining how the computation suits the problem situation ("five hours times ninety miles per hour."). Hence, the failure to explain the inappropriate computation is the creativity dampener that sharply reduces micro-creativity. As this example shows, the breakpoint only identifies the beginning of a period of sharply lower (or higher) micro-creativity; it does not necessarily identify the precise moment of the creativity dampener (or ignitor).

A group member's action sharply increased micro-creativity (*creativity ignitors;* e.g., insight about the problem, drawing a diagram, etc.) at 17 of the breakpoints (11, 6). After the period of silence in the previous segment, a creativity ignitor follows in line 44.

44 Jan: But only the ship goes five hours. The helicopter don't need five hours to catch the ship.
45 Xia: One and a half hours? (1 second of silence) Times ninety?
46 Ron: [Presses calculator keys 1.5 × 90 =] one thirty-five.
47 Jan: Less than an hour and a half.
48 Xia: Uh-huh. Maybe an hour and a quarter.

In line 44, Jan identifies the problem with Ron's computation ("But only the ship goes five hours. The helicopter don't need five hours to catch the ship.") and ignites a series of new ideas, starting with Xia's suggestion that the helicopter travels "one and a half hours."

The 10 successful groups that correctly solved the problem showed higher micro-creativity than did the 10 unsuccessful groups. However, they did not do so consistently, as micro-creativity varied more across time periods within a group than across groups. Less than 0.1% of the micro-creativity variance occurred at the classroom or group level (M = 0.000, SE = 0.001), 79% was across time periods (M = 3.46, SE = 0.68) while 21% was across speaker turns (M = 0.92, SE = 0.02). As neither the micro-creativity variance across classrooms nor that across groups was significant, a 2-level model (time periods and turns) was sufficient.

4.2. Effects on micro-creativity

Disagreements, polite actions, mathematics grade, and social status were all linked to greater micro-creativity. In contrast, diversity, questions, and status variances were not linked to micro-creativity. So, the diversity hypothesis (H-1), status difference hypotheses (H-4), and part of the argumentation hypothesis (H-2, group gap questions) were not supported.

4.2.1. Agreements vs. disagreements

Compared to agreements, disagreements yielded greater micro-creativity, supporting the first half of the argumentation hypothesis (H-2, see Table 4 and Fig. 2). Agreement yielded less micro-creativity, possibly because group members might have been overly inclined to agree with another (even with wrong ideas) due to social motives or face concerns.

Disagreements often yielded higher micro-creativity both when group members recognized flaws in wrong ideas and when they did not. Following a flawed idea, the next speaker was more likely to rudely disagree (+5%) and less likely to agree (−17%). Specifically, rude disagreements that identified flawed ideas in the previous speaker turn (−1) or two speaker turns ago (−2) increased micro-creativity (+9% and +60%). Without a disagreement, wrong ideas often yielded more wrong ideas (+11%), showing the importance of identifying flawed ideas via disagreement.

Disagreements that did not follow wrong ideas also yielded higher micro-creativity compared to agreement, especially rude disagreement by the previous speaker (+10%) or polite disagreement by the current speaker (+11%). These results are consistent with claims that disagreements might increase micro-creativity not only through identifying errors, but also through increased attention, legitimacy of different views, or creative misinterpretations (De Dreu & West, 2001; Nemeth & Chiles, 1988; Nemeth & Rogers, 1996).

4.2.2. Polite vs. rude actions

Compared to a polite disagreement, a person who disagrees rudely shows less micro-creativity, but can increase group members' micro-creativity. Disagreeing rudely rather than politely reduces one's own micro-creativity, exceeding the smaller effect of increasing others' micro-creativity (−15% > +10%; see Table 4, rows 2 and 4). The negative effect of rudely disagreeing is especially strong in groups that did not solve the problem (−26%; Table 4, row 6). In the special case when the previous speaker disagrees rudely with a wrong idea, it increases the current speaker's micro-creativity (partial effect of rudely disagree [−1] * Wrong [−2] = +60%; Table 4, row 5) (After a wrong idea, a person who disagrees rudely rather than politely shows less micro-creativity, −15% > +9%; Table 4, rows 2 and 3). In short, disagreeing rudely rather than politely generally reduces overall micro-creativity, except in the special case of disagreeing rudely with a wrong idea by the previous speaker. Consider the following chain reaction of rude disagreements.

33 Jan: That don't make sense.
34 Ron: Yes, it does.
35 Jan: No, it doesn't.

After Jan rudely disagrees with Ron's last computation (''that doesn't make sense''), Ron reflexively disagrees (''yes, it does''), and Jan repeats her disagreement (''no, it doesn't''). During these chains of

Table 4
Total effects of each explanatory variable on target outcome and explanatory variables

	Explanatory variable (E)	Target (T)	P(T | E)[a] (%)	P(T|~ E)[b] (%)	Effect (%)
(1)	Agree[c]	Micro-creativity	25	36	−11
(2)	Rude disagree[c]	Micro-creativity	17	31	−15
(3)	Rude disagree[c] * Wrong (−1)	Micro-creativity	39	30	+9
(4)	Rude disagree[c] (−1)	Micro-creativity	40	30	+10
(5)	Rude disagree[c] (−1) * Wrong (−2)	Micro-creativity	90	30	+60
(6)	Rude disagree[c] (−1) * Unsolved	Micro-creativity	6	32	−26
(7)	Rude disagree[c] (−1)	Agree	35	58	−22
(8)	Rude disagree[c] (−1) * Unsolved	Rude disagree	26	8	+18
(9)	Wrong (−1)	Micro-creativity	37	29	+8
(10)	Wrong (−2)	Micro-creativity	19	32	−13
(11)	Wrong (−1)	Agree	41	58	−17
(12)	Wrong (−1)	Rude disagree	14	9	+5
(13)	Wrong (−2)	Rude disagree (−1)	14	9	+5
(14)	Wrong (−2)	Wrong (−1)	24	12	+11
(15)	Command (−1)	Micro-creativity	22	31	−9
	A 10% increase above the mean				
(16)	Social status	Micro-creativity			+3
(17)	Social status (−1)	Micro-creativity			+2
(18)	Mathematics grade (−1)	Micro-creativity			+2

[a] Probability that the target occurs, given that the explanatory variable does occur.
[b] Probability that the target occurs, given that the explanatory variable does not occur.
[c] The baseline value for comparison is polite disagreement. See Kennedy (2004) for details.

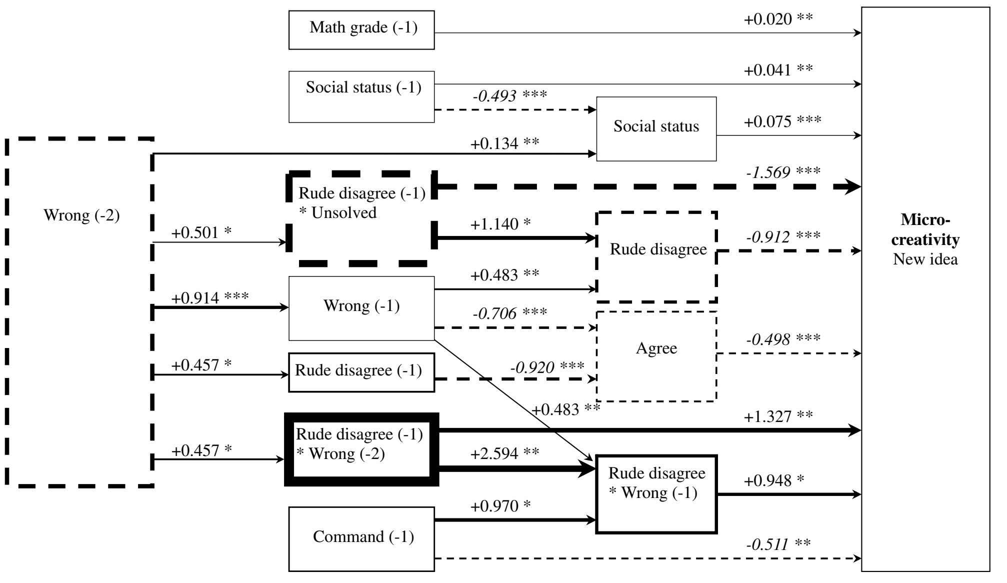

Fig. 2. Path analysis of significant explanatory variables of micro-creativity using multilevel Logit. Values are standardized parameter coefficients. Explanatory variables inside solid boxes indicate positive overall effects, while those inside dashed boxes indicate negative overall effects. Solid arrows (→) indicate positive direct effects, while dashed arrows (– →) indicate negative direct effects. Line widths indicate effect sizes.

rude disagreements, Jan and Ron re-asserted their positions instead of creating new ideas. In contrast, students often addressed polite disagreements with new ideas.

71 Ana: A hundred miles.
72 Juan: I don't understand how we got a hundred miles
73 Ana: Five times twenty-two, oh, [presses calculator keys 5 $\times$ 22=] a hundred and ten.

Juan disagrees politely by stating his inadequacy ("I don't understand,"), shared positioning ("we"), and shared understanding ("a hundred miles"). In response, Ana addresses Juan's concern by explaining her computation ("five times twenty-two"). Suddenly, she recognizes her error ("oh"), and corrects it with a new idea ([presses calculator keys 5 $\times$ 22=] "a hundred and ten").

If the previous speaker rudely disagrees (-1) with a wrong idea (-2) however, group micro-creativity increases sharply (+60%). Although students who rudely disagreed showed less micro-creativity, they identified flaws to help their group members create new ideas, thereby increasing group micro-creativity. Rude disagreements also yielded fewer agreements by the next speaker (-22%). In unsuccessful groups, rude disagreements yielded more subsequent rude disagreements (+18%).

Commands (-1) yielded less micro-creativity (-9%) than did questions or statements. However, commands (-1) that included wrong ideas (-1) were slightly more easily recognized by other group members who rudely disagreed with it to create a new idea (+1%).

Students with higher mathematics grades or social statuses enhanced their group's micro-creativity in two ways. Higher social status students showed greater micro-creativity (+3% when their social status exceeded the mean by 10%). Also, speakers with higher mathematics grades or social statuses enhanced other students' micro-creativity (+2% when the previous speaker's mathematics grade or social status exceeded the mean by 10%) (the negative relationship between social status (-1) and social status was likely an artifact of the limited number of group members. After a higher social status student spoke, the remaining students who can speak were likely of lower social status). Note that these individual characteristics have much weaker links to micro-creativity, compared to group processes.

No other variables showed significant effects. Aside from the different effects of rude disagreements across groups, other explanatory variables showed no significant differences across groups or across time periods. Hence, the remaining explanatory variable effects are candidates for broader (possibly universal) effects across time periods and across groups.

This model had a 70% accuracy rate for predicting micro-creativity in any given turn (y_{ijk}^* vs. y_{ijk}). Furthermore, the final model's Q-statistics showed no significant serial correlation of residuals in any of the 20 groups. So, the time-series model was likely appropriate. Unlike the coarse, non-dynamic analysis result that 76% of the group level variance was explained (see Appendix A), only 6% of the speaker turn variance was explained when DMA was used, suggesting that many other omitted variables should be added to this model to provide a fuller account of the micro-contexts that influence group micro-creativity.

5. Discussion

The explosion of information and specializations are increasingly transforming creativity into a group endeavor rather than an individual one (Paulus & Brown, 2003). In this study, a step is taken toward understanding group creativity by analyzing how group processes (such as argumentation) affect the creation of new ideas (micro-creativity). As past research shows that group processes affect group outcomes, these same group processes might affect micro-creativity, which might affect problem solving success. Analyses of 20 groups of students' algebra problem solving showed that groups with more micro-creativity often solved the problem successfully. Also, statistically-identified watershed breakpoints separated high micro-creativity time periods from low micro-creativity time periods. Across all groups and time periods, disagreements and politer actions increased micro-creativity.

5.1. Successful groups: More micro-creativity, smaller rude disagreement effect

In this study, groups with higher mathematics grades, greater micro-creativity, or a weaker rude disagreement effect on micro-creativity were more likely to solve the problem successfully. As expected, groups with higher achieving students were more successful. Also, groups showing higher micro-creativity were more successful, consistent with earlier studies (e.g., Paulus & Brown, 2003).

Meanwhile, rude disagreements had a stronger effect in unsuccessful groups, yielding 26% less micro-creativity than in successful groups. Groups that tolerated rude disagreements were more likely to build on their criticisms to create new ideas and successfully solve the problem. Thus, rude disagreement was moderated by an unknown group variable linked to the success of the group.

5.2. Breakpoints separate high and low micro-creativity time periods

Dynamic multilevel analysis (DMA) was used to identify watersheds (breakpoints) that radically changed group problem solving processes and separated high micro-creativity time periods from low micro-creativity time periods. An exploratory classification of the 53 breakpoints yielded three broad categories: off-task ↔ on-task transitions, creativity ignitors, and creativity dampeners.

Breakpoint characteristics did not differ significantly across successful and unsuccessful groups. Furthermore, no independent variable showed significantly different effects across time periods.

5.3. Disagreements and politer actions increase micro-creativity

Compared to its a priori characteristics, a group's recent actions (micro-time context) were more strongly linked to micro-creativity in this study. Thus, inferring group processes from group member characteristics can be insufficient for understanding the relationships among group processes, which can require examination of actual actions and interactions. Specifically, both disagreements and politer actions were linked to greater micro-creativity.

5.3.1. Agreements vs. disagreements

Unlike agreements, disagreements (with incorrect or correct ideas) yielded 11% greater micro-creativity. Past studies have shown that social motives and face concerns overly incline people to agree with one another (Burgoon et al., 1993; Person et al., 1995; Tudge, 1989). Likewise, group members in this study were overly inclined to agree with one another (even with wrong ideas) and were less likely to create their own ideas, thereby yielding less micro-creativity.

In contrast, disagreements often yielded higher micro-creativity both when group members recognized flaws in wrong ideas (+60%) and when they did not (+10%). After a wrong idea, group members were more likely to disagree (+22%) and to disagree rudely (+5%), suggesting that they often recognized flaws while solving this problem. After a disagreement with a wrong idea, a student was 60% more likely to create a new idea, possibly from the useful part of the flawed idea. Thus, wrong ideas might have served as kindling for micro-creativity. This benefit often outweighed the danger of accepting wrong ideas while solving this problem.

Furthermore, disagreements with correct ideas also increased micro-creativity by +10%, consistent with studies showing that disagreements can also aid problem solving through increased attention, legitimacy of different views, or creative misinterpretations. For example, disagreements might stimulate the attention of group members, helping them consider more aspects of the situation from more perspectives to increase micro-creativity indirectly (De Dreu & West, 2001). By legitimizing different opinions, disagreements can encourage group members to express their ideas, also indirectly aiding micro-creativity (Nemeth & Chiles, 1988). Lastly, disagreements with correct ideas might yield creative misinterpretations that directly increase micro-creativity (Chiu, 1997).

5.3.2. Polite vs. rude actions

Politer actions often yielded greater micro-creativity than did ruder actions (rude disagreements, commands). Overall, polite disagreements yielded 15% greater micro-creativity compared to rude disagreements. In one type of situation however, a group member's rude disagreement with a wrong idea reduced his or her own micro-creativity but increased group micro-creativity 60%; in this case, a rude disagreement highlights the flawed ideas, thereby helping other group members create new ideas. Meanwhile, commands yielded 9% less micro-creativity than did questions or statements.

In general, rude behaviors can harm both a group's social relations and its cognitive problem solving (Chiu & Khoo, in press). In contrast, politer disagreements arguably support social relations, reduce the likelihood of emotional retaliation, encourage understanding of new ideas, and facilitate micro-creativity (Chiu & Khoo, 2003). Likewise, questions and statements are politer than commands. Unlike commands, questions and statements allow target listeners greater freedom to pursue their ideas and increase micro-creativity (Brown & Levinson, 1987; Chiu & Khoo, 2003; Tsui, 1992). Like psychological demands (such as mathematics anxiety, Ashcraft & Krause, 2007), rude behaviors might create additional social demands that occupy working memory, reduce available working memory for problem solving, and thereby reduce performance, in this case, micro-creativity. Future studies can test these causal mechanisms.

5.4. Limitations

This study's limitations include its small sample of classrooms and groups, limited problem content, single setting, and methodological assumptions. Due to the small number of teachers, classrooms, and groups, the data were not necessarily representative of group interactions in classrooms across different subject domains. Likewise, results might differ across schools or across countries. In particular, these students were not accustomed to working together in groups, so students with substantial group preparation or collaboration experience might behave differently. Group preparation (e.g., scripts or roles) have shown mixed results (e.g., Strijbos, Martens, Jochems, & Broers, 2007; Webb & Farivar, 1999; Weinberger, Ertl, Fischer, & Mandl, 2005), possibly due to differences in implementation. Effective group preparation might increase micro-creativity by increasing desirable group processes (e.g., polite disagreement) and by reducing undesirable group processes (e.g., rude disagreements). Still, the influence of each group process on micro-creativity might be similar to the results in this study. Future research can show if these effects differ with effective group preparation, possibly because the rarity of an undesirable action magnifies its impact when it does occur.

DMA has two assumptions and requires a minimum sample size. Like other regressions, a linear combination of explanatory variables and independent, identically distributed residuals are both assumed for DMA (Non-linear functions can be modeled as individual or multiple explanatory variables, e.g., age^2). DMA also has modest sample size requirements. Green (1991) proposed the following heuristic sample size, N, for a multiple regression with M explanatory variables and an expected explained variance R^2 of the outcome variable:

$$N > 8 * (1 - R^2)/R^2 + M - 1 \quad (3)$$

For a large model of 25 explanatory variables with a small expected R^2 of 0.1, the required sample size is 96 speaker turns (= 8 * (1 − 0.10)/0.10 + 25 − 1). The required sample size is smaller for models with fewer explanatory variables or with a larger expected R^2. In practice, two groups of students talking for half an hour will often yield over 100 speaker turns, sufficient for DMA. Thus, researchers can analyze seemingly "qualitative" data sets through both qualitative and quantitative methods (e.g., both traditional discourse analysis and DMA's statistical discourse analysis).

5.5. Implications for researchers, teachers, and students

In this study, the micro-context of recent group processes (such as argumentation) influenced group micro-creativity, as shown by a statistical discourse analysis. Due to the small number of groups, the data were not necessarily representative of group interactions. If future studies show similar findings, these results have the following implications for researchers, students, and teachers.

There are two major implications for researchers. First, the micro-time context (breakpoints, recent actions) affects group micro-creativity. Specifically, watershed breakpoints and sequences of actions and interactions by the three most recent speakers affected the current speaker's micro-creativity. In addition to modeling an activity's broader macro-context, researchers might develop better understanding of student argumentation and collaborative group processes by modeling the micro-context of time and the differences in group processes across these micro-contexts.

Second, a new statistical discourse analysis method, dynamic multi-level analysis (DMA) was applied (Chiu & Khoo, 2005). Using DMA, watershed breakpoints were statistically identified, and individual actions and social interactions over time were modeled. With the breakpoint method, watersheds that radically changed the nature of a group's problem solving were identified. An exploratory analysis suggests three types of breakpoints: on-task ↔ off-task transitions, creativity ignitors (e.g., insights), and creativity dampeners (e.g., inadequate explanations). Meanwhile, the use of relational variables across speaker turns (e.g., disagreements), multi-level Logit/Probit, lag variables, and serial correlation tests modeled social interactions within a micro-context of time along with both group and time period differences. By identifying differences in effects across groups or across time periods, DMA can show moderator effects of unexamined variables at the group or time period levels. Only rude disagreements showed different effects across groups, as rude disagreement was moderated by an unknown group variable linked to the success of the group.

The results of this study have practical implications as well. Teachers can encourage students to evaluate one another's ideas carefully, speak politely, and avoid unproductive responses to rude actions. Teachers can ask students to consider new ideas carefully, discourage impulsive social confirmations, and support polite disagreements that identify specific flaws. Furthermore, teachers can model and reinforce respectful and polite behaviors. When students behave rudely, teachers can respond calmly by listening to the content while addressing the rude behavior. Then, students can capitalize on the ideas within the rude disagreement, avoid impulsive retaliations, and increase their micro-creativity.

6. Conclusion

This study of eighty high school students' group problem solving showed that the micro-time context (created by watershed breakpoints and recent group member actions) was linked to group micro-creativity (new ideas). Three types of watershed breakpoints (creative ignitors, creative dampeners, and on-task ↔ off-task transitions) radically changed the groups' problem solving processes and separated high micro-creativity time periods from low micro-creativity time periods. Furthermore, recent

Table A1
An inferior non-dynamic analysis predicting percentage of new ideas in each group showing regression coefficients (with standard errors)

Variable	Predict percentage of new ideas
% Wrong idea	0.90** (0.29)
% Rude disagreement	−2.34*** (0.45)
% Command	−1.45** (0.40)
Constant	0.55*** (0.05)
R-squared	0.76

A non-dynamic analysis yields the following results. Similar to the DMA results, percentage of new ideas is positively correlated with percentage of wrong ideas and negatively correlated with rude disagreements and commands. In this coarser analysis however, the effects of mathematics grade, social status, agreement, and several interaction effects are not significant, unlike the DMA results.

Furthermore, a non-dynamic analysis does not identify which earlier actions are linked to the likelihood of a new idea at each speaker turn. The non-dynamic analysis also yields a deceptively high explained variance at only the group level, ignoring the much larger variance at the time period and speaker turn levels. Lastly, a non-dynamic analysis does not identify watershed breakpoints that separate the problem solving session into distinct time periods.

Hence, a dynamic multilevel analysis is superior to a non-dynamic analysis.

group members' disagreements yielded greater micro-creativity, compared to agreements. Rude actions (rude disagreements and commands) often yielded less micro-creativity than polite actions (polite disagreements and questions/statements), with one exception; after a wrong idea, a rude disagreement increased group members' micro-creativity more than did a polite disagreement. Together, these results show how the group micro-time context can affect a person's micro-creativity.

Appendix A. A non-dynamic analysis of the same data

See Table A1.

References

Ashcraft, M. H., & Krause, J. A. (2007). Working memory, math performance, and math anxiety. *Psychonomic Bulletin & Review, 14*, 243–248.

Bales, R. F. (2001). *Social interaction systems*. New Brunswick, NJ: Transaction Publishers.

Barnes, D., & Todd, F. (1995). *Communication and learning revisited*. Portsmouth, NH: Boynton/Cook.

Barron, B. (2003). When smart groups fail. *The Journal of the Learning Sciences, 12*(3), 307–359.

Benjamini, Y., Krieger, A. M., & Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika, 93*, 491–507.

Berger, J., Cohen, B. P., & Zelditch, M. (1972). Status characteristics and social interaction. *American Sociological Review, 37*, 241–256.

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. New York: Cambridge University Press.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models*. Newbury Park: Sage.

Buchs, C., Butera, F., Mugny, G., & Darnon, C. (2004). Conflict elaboration and cognitive outcomes. *Theory Into Practice, 43*, 23–30.

Burgoon, J. K., Dillman, L., & Stern, L. (1993). Adaptation in dyadic interaction. *Communication Theory, 3*, 295–316.

California Department of Education (2005). *Academic performance index*. Retrieved on May 8, 2007, from <http://www.cde.ca.gov/ta/ac/ar/index.asp>.

Chaiken, S., & Stangor, C. (1987). Attitudes and attitude change. *Annual Review of Psychology, 38*, 575–630.

Chiu, M. M. (1997). *Building on diversity*. Los Angeles: University of California, Los Angeles (ERIC Documentation Reproduction Service No. ED 410 325).

Chiu, M. M. (2000a). Group problem solving processes: Social interactions and individual turns. *Journal for the Theory of Social Behavior, 30*(1), 27–50.

Chiu, M. M. (2000b). Status effects on solutions, leadership, and evaluations during group problem solving. *Sociology of Education, 73*(3), 175–195.

Chiu, M. M. (2001). Analyzing group work processes. In F. Columbus (Ed.). *Advances in psychology research* (Vol. 4, pp. 193–222). Huntington, NY, US: Nova Science.

Chiu, M. M., & Khoo, L. (2003). Rudeness and status effects during group problem solving. *Journal of Educational Psychology, 95*, 506–523.

Chiu, M. M., & Khoo, L. (2005). A new method for analyzing sequential processes: Dynamic multi-level analysis. *Small Group Research, 36*, 600–631.

Chiu, M. M., & Khoo, L. (in press). Conflict between cooperative and competitive goals during group problem solving. *Human Communication Research*.

Civettini, N. H. W. (2007). Similarity and group performance. *Social Psychology Quarterly, 70*, 262–271.
Cobb, P. (1995). Mathematical learning and small-group interaction: Four case studies. In P. Cobb & H. Bauersfeld (Eds.), *The emergence of mathematical meaning* (pp. 25–130). Hillsdale, NJ: Erlbaum.
Cohen, E. G. (1982). Expectation states and interracial interaction in school settings. *Annual Review of Sociology, 8*, 109–235.
Cohen, E. G. (1994). Restructuring the classroom. *Review of Educational Research, 64*, 1–35.
De Dreu, C. K. W., & West, M. A. (2001). Minority dissent and team innovation. *Journal of Applied Psychology, 86*, 1191–1201.
De Lisi, R., & Goldbeck, S. L. (1999). Implications of Piagetian theory for peer learning. In A. M. O'Donnell & A. King's (Eds.), *Cognitive perspectives on peer learning* (pp. 3–37). Mahwah, NJ: Erlbaum.
Dembo, M. H., & McAuliffe, T. J. (1987). Effects of perceived achievement and grade status on social interaction and influence in cooperative groups. *Journal of Educational Psychology, 79*, 415–423.
Ding, C. S., Kim, S., & Richardson, L. I. (2007). Do mathematical gender differences continue? *Educational Studies, 40*, 279–295.
Doise, W., Mugny, G., & Perret-Clermont, A. N. (1975). Social interaction and development of cognitive operations. *European Journal of Social Psychology, 5*, 367–383.
Dunbar, K. (1997). How scientists think. In T. B. Ward, S. M. Smith, & J. Vaid (Eds.), *Creative Thought* (pp. 461–493). Washington, DC: APA.
Fleming, S., Shire, J., Jones, D., Pill, A., & McNamee, M. (2004). Continuing professional development. *Journal of Further and Higher Education, 28*, 165–177.
Gersick, C. J. G. (1988). Time and transition in work teams: Toward a new model of group development. *Academy of Management Journal, 31*, 9–41.
Goldstein, H. (1995). *Multilevel statistical models*. Sydney: Edward Arnold.
Goldstein, H., Healy, M. J. R., & Rasbash, J. (1994). Multi-level time series models with applications to repeated measures data. *Statistics in Medicine, 13*, 1643–1655.
Green, S. B. (1991). How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research, 26*, 499–510.
Gruber, H. E., & Wallace, D. B. (1999). The case study method and evolving systems approach for understanding uniquely creative people at work. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 93–115). New York: Cambridge University Press.
Hackman, M. Z., & Johnson, C. E. (2000). *Leadership*. Prospect Heights, IL: Waveland.
Holtgraves, T. (1997). YES, BUT ... Positive politeness in conversation arguments. *Journal of Language and Social Psychology, 16*, 222–239.
Janis, I. L. (1989). *Crucial decisions*. New York: Free Press.
Jehn, K. A., & Mannix, E. A. (2001). The dynamic nature of conflict. *Academy of Management Journal, 44*, 238–251.
Johnson, D. W., & Johnson, R. T. (1994). *Learning together and alone: Cooperative, competitive, and individualistic learning* (4th ed.). Boston: Allyn & Bacon.
Johnson, D., Johnson, R., & Smith, K. (2007). The state of cooperative learning in postsecondary and professional settings. *Educational Psychology Review, 19*, 15–29.
Kennedy, P. (2004). *Guide to econometrics*. Cambridge: Blackwell.
Krippendorff, K. (2004). *Content analysis*. Thousand Oaks, CA: Sage.
Kuhn, D., Shaw, V., & Felton, M. (1997). Effects of dyadic interaction on argumentative reasoning. *Cognition and Instruction, 15*, 287–315.
Larson, J. R. Jr., (2007). Deep diversity and strong synergy. *Small Group Research, 38*, 413–436.
Larson, J. R., Jr., Christiansen, C., Abbott, A. S., & Franz, T. M. (1996). Diagnosing groups. *Journal of Personality and Social Psychology, 71*, 315–330.
Laughlin, P. R., Hatch, E. C., Silver, J. S., & Boh, L. (2006). Groups perform better than the best individuals on letters-to-numbers problems. *Journal of Personality and Social Psychology, 90*, 644–651.
Ljung, G., & Box, G. (1979). On a measure of lack of fit in time series models. *Biometrika, 66*, 265–270.
Ma, X., & Xu, J. (2004). The causal ordering of mathematics anxiety and mathematics achievement. *Journal of Adolescence, 27*, 165–179.
McLeod, P. L., Lobel, S. A., & Cox, T. H. Jr., (1996). Ethnic diversity and creativity in small groups. *Small Group Research, 27*, 2, 248–26.
Milliken, F. J., Bartel, C. A., & Kurtzberg, T. R. (2003). Diversity and creativity in work groups. In P. B. Paulus & B. A. Nijstad's (Eds.), *Group creativity* (pp. 32–62). Oxford: Oxford.
Mulryan, C. M. (1992). Student passivity during cooperative small groups in mathematics. *The Journal of Educational Research, 85*, 261–273.
Mumford, M. D., & Gustafson, S. G. (1988). Creativity syndrome: Integration, application, and innovation. *Psychological Bulletin, 103*, 27–43.
Nemeth, C. J., & Chiles, C. (1988). Modeling courage: The role of dissent in fostering independence. *European Journal of Psychology, 18*, 275–280.
Nemeth, C. J., & Rogers, J. (1996). Dissent and the search for information. *British Journal of Social Psychology, 35*, 67–76.
Nijstad, B. A., Diehl, M., & Stroebe, W. (2003). Cognitive stimulation and interference in idea generating groups. In P. B. Paulus & B. A. Nijstad (Eds.), *Group creativity: Innovation through collaboration* (pp. 137–159). New York: Oxford University Press.
Orlitzky, M., & Hirokawa, R. Y. (2001). To err is human, to correct for it divine. *Small Group Research, 32*, 313–341.
Paulus, P. B. (2000). Groups, teams, and creativity. *Applied Psychology, 49*, 237–262.
Paulus, P. B., & Brown, V. (2003). Ideational creativity in groups. In P. B. Paulus & B. A. Nijstad (Eds.), *Group creativity* (pp. 110–136). New York: Oxford University Press.
Person, N. K., Kreuz, R. J., Zwaan, R., & Graesser, A. C. (1995). Pragmatics and pedagogy. *Cognition and Instruction, 13*, 161–188.
Piaget, J. (1985). The equilibrium of cognitive structures. In T. Brown & K. L. Thampy (Eds.), *Trans*. Chicago: University of Chicago Press.
Pirie, S. (1991). Peer discussion in the context of mathematical problem solving. In D. Durkin & B. Shire (Eds.), *Language in mathematical education: Research and practice* (pp. 143–161). Buckingham: Open University Press.
Rasbash, J., & Woodhouse, G. (1995). *MLn command reference*. London: Multilevel Models Project, Institute of Education.
Sfard, A. (2007). When the rules of discourse change, but nobody tells you. *Journal of the Learning Sciences, 16*, 565–613.

Sinclair, J., & Coulthard, M. (1992). Towards an analysis of discourse. In M. Coulthard (Ed.), *Advances in spoken discourse analysis* (pp. 1–34). London: Routledge.
Stasser, G., & Birchmeier, Z. (2003). Group creativity and collective choice. In P. B. Paulus & B. A. Nijstad's (Eds.), *Group creativity* (pp. 85–109). Oxford: Oxford University.
Stasser, G., & Taylor, L. (1991). Speaking turns in face-to-face discussions. *Journal of Personality and Social Psychology, 60*, 675–684.
Stasson, M. F., & Bradshaw, S. D. (1995). Explanation of individual-group performance differences. *Small Group Research, 26*, 296–308.
Sternberg, R. J., & Lubart, T. I. (1999). The concept of creativity. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 3–15). New York: Cambridge University Press.
Strijbos, J-W., Martens, R. L., Jochems, W. M. G., & Broers, N. J. (2007). The effect of functional roles on perceived group efficiency during computer-supported collaborative learning. *Computers in Human Behavior, 23*, 353–380.
Swann, W. B. Jr., (2005). The self and identity negotiation. *Interaction Studies, 6*, 69–83.
Swann, W. B., Jr., Kwan, V. S. Y., Polzer, J. T., & Milton, L. P. (2003). Fostering group identification and creativity in diverse groups. *Personality and Social Psychology Bulletin, 29*, 1396–1406.
Teasley, S. D., & Roschelle, J. (1993). Constructing a joint problem space. In S. P. Lajoie & S. J. Derry (Eds.), *Computers as cognitive tools* (pp. 229–258). Mahwah, NJ: Erlbaum.
Tracy, K., & Tracy, S. J. (1998). Rudeness at 911: Reconceptualizing face and face attack. *Human Communication Research, 25*, 225–251.
Tsui, A. (1992). A functional description of questions. In M. Coulthard (Ed.), *Advances in spoken discourse analysis* (pp. 89–110). London: Routledge.
Tudge, J. R. H. (1989). When collaboration leads to regression: Some negative consequences of socio-cognitive conflict. *European Journal of Social Psychology, 19*, 123–138.
Turner, J. C. (2005). Explaining the nature of power: A three-process theory. *European Journal of Social Psychology, 35*, 1–22.
Turner, J. C., & Horvitz, T. (2001). The dilemma of threat. In M. E. Turner (Ed.), *Groups at work: Theory and research* (pp. 445–470). Mahwah, NJ: Erlbaum.
Webb, N. M. (1984). Sex differences in interaction and achievement in cooperative small groups. *Journal of Educational Psychology, 76*, 33–44.
Webb, N. M., & Farivar, S. (1999). Developing productive group interaction in middle school mathematics. In A. M. O'Donnell & A. King's (Eds.), *Cognitive perspectives on peer learning* (pp. 117–149). Mahwah, NJ: Erlbaum.
Weinberger, A., Ertl, B., Fischer, F., & Mandl, H. (2005). Epistemic and social scripts in computer-supported collaborative learning. *Instructional Science, 33*, 1–30.
Wentzel, K. R., & Watkins, D. E. (2002). Peer relationships and collaborative learning as contexts for academic enablers. *School Psychology Review, 31*(3), 366–377.
West, M. A. (2000). Reflexivity, revolution, and innovation in work teams. *Advances in Interdisciplinary Studies of Work Teams, 5*, 1–29.
Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27*, 458–477.