Effects of social metacognition on micro-creativity: Statistical discourse analyses of group problem solving

Ming Ming Chiu, Department of Educational Psychology, 314 Ho Tim Building,
CUHK, Shatin, NT, Hong Kong, mingmingchiu@cuhk.edu.hk

Abstract: This study examines how social metacognition (including evaluations and questions) affected micro-creativity during group problem solving. Twenty groups of high school students were videotaped as they solved a mathematics problem. Analyses of the 2,951 conversation turns showed that the likelihood of a correct contribution (CC, a measure of micro-creativity) was higher after a group member expressed a wrong, new idea, correctly evaluated an idea, or justified an idea. In contrast, the likelihood of a CC was lower after a group member disagreed rudely or agreed. Meanwhile, group-level properties (racial diversity, gender diversity, and degree of status differences) did not significantly affect the likelihood of a CC. A CC was more likely after a justification in successful groups than in unsuccessful groups. CCs did not occur uniformly, as some time periods had many CCs while others had few CCs. Furthermore, agreements and correct evaluations had different effects across time periods within a group.

Past research on the development of original ideas that are useful or influential (*creativity*) has largely focused on individuals (Sternberg & Lubart, 1999). However, the explosion of information and specializations will increasingly require teams with diverse skills and knowledge to create innovations (*group creativity*; Sawyer, 2004). Still, researchers have not systematically examined how group processes affect creative moments (*group micro-creativity*). This study takes a step in this direction by analyzing how micro-creativity occurs in twenty groups of students solving an algebra problem. By understanding how group processes affect micro-creativity, group members can work together more creatively.

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 (Sternberg & Lubart, 1999). Furthermore, micro-creativity is measured via a correct, new idea (*correct contribution*, CC). Hence, the CC's "new" component is relative to the group members' experiences. The CC's "correct" component refers to the intersection of the problem situation and the relevant field's principles (in this case, algebra). Thus, a CC is a new idea for the group members that is consistent with both the given problem situation and algebra principles.

I modeled group, student, and speaker turn characteristics that helped or hindered micro-creativity during group problem solving in a high school algebra class. This study contributes to the research literature in four ways. First, I showed when CCs occur, whether they occur uniformly through a problem-solving session or more frequently in some time periods than in others. Second, this study showed how the local time context created by prior speakers' actions (CCs, justifications, etc.) and interactions affected the likelihood of creating a CC. Third, I tested whether the above effects differed across groups or across time periods. Lastly, I tested these hypotheses with a new statistical discourse analysis method, dynamic multilevel analysis (Chiu & Khoo, 2005).

Group processes and micro-creativity

Diverse views and social metacognition can aid CC creation (Chiu, in press; Paulus & Brown, 2003). Meanwhile, public self-image (face) and status concerns can hinder creation of CCs (Brown & Levinson, 1987).

Group problem solving processes that aid micro-creativity

Diversity

Groups with diverse views can create more ideas. Specifically, group diversity in nationality or ethnicity can increase the number of perspectives, number of ideas, and quality of ideas (McLeod, Lobel, & Cox, 1996). Furthermore, groups that create many ideas, representations, and solution proposals, are more likely to find a correct or optimal solution (Paulus & Brown, 2003).

Group members with diverse views can also build on idiosyncratic ideas to create CCs through processes such as sparked ideas, jigsaw pieces, and creative misinterpretations (Paulus & Brown, 2003; Chiu, 1997). 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 CC. Like fitting jigsaw pieces together, group members also can put together different pieces to create a CC. Lastly, a person might also misinterpret another person's incorrect idea to create a CC. Thus, even wrong, new ideas can lead to CCs.

Group members' diverse views can help identify flaws and refine incorrect ideas. (Orlitzky & Hirokawa, 2001). By creating more ideas (including wrong ideas) and evaluating their validity correctly, groups are more likely to create CCs. Hence, group members need not only build on correct ideas.

Social Metacognition

Social metacognition. Successful group problem solving often involves social metacognition and justifications in the cognitive/problem content space (Roschelle, 1992; Teasley & Roschelle, 1993). Whereas individual metacognition is monitoring and controlling one's own knowledge (Hacker & Bol, 2004), emotions, and actions, *social metacognition* is *group members'* monitoring and control of *one another's* knowledge, emotions, and actions. Students use social metacognitive strategies to evaluate one another's ideas and recognize problems. In response to other group members' monitoring, people often justify their own views to convince others (Amason, 1996; Cobb, 1995).

According to the *functional theory of group decision-making*, group members at least implicitly evaluate the previous speaker's action and problem solving approach (Orlitzky & Hirokawa, 2001). For example, Jay says "five times six is eleven." Kay 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 this trajectory (Chiu, 2000).

Evaluations can also be right or wrong in some contexts (e.g., algebra problems). Correct evaluations support correct ideas ("yes, five times six is thirty,") or identify flawed ideas ("No! Five times six is not eleven,"), and thus, can help create a shared knowledge base that group members might use to build new CCs. In contrast, incorrect evaluations reject CCs ("wrong, five times six is NOT thirty,") or accept flawed ideas ("right, five times six is seven,"), embedding flaws in their shared knowledge base. Group members using this shared knowledge base might bring these flaws into their new ideas, resulting in wrong ideas rather than CCs. A group's collective attention and diverse perspectives can help it evaluate ideas correctly and create a shared knowledge base to aid creation of CCs (Chiu, 2000).

When group members recognize problems or difficulties (perturbations), they can disagree or ask a question (Piaget, 1985). Piaget (1985) defined two types of perturbations: (a) obstacles, which give negative feedback and (b) lacunae, gaps in understanding. Disagreements indicate obstacles to be overcome (e.g., "no, that's wrong, five times six isn't eleven"). Thus, disagreements can show the need for CCs and identify flaws to avoid, thereby motivating and aiding the creation of CCs.

Meanwhile, a question (e.g., Jay asks, "what's five times eight?") might indicate an individual gap or a group gap. For an individual gap question, other group members who know the answer can help her (e.g., Kay says "forty"). Thus, individual gap questions encourage review of old ideas rather than the creation of CCs. In contrast, no one knows the answer to a group gap question, which motivates the need for a CC and points to a new direction for creating one. Group members' diverse views and levels of knowledge facilitate both perturbations and responses to them. In short, these perturbations can motivate and inform creation of CCs.

After perturbations provoke new ideas, group members often justify themselves. Chiu and Khoo (2003) showed that members of successful groups often anticipated criticisms and justified their new ideas. Likewise, after a person disagrees with a proposal (e.g. Kay), the original proposer (Jay) might try to justify it by linking it to data, using a warrant, or supporting a warrant with backing. In response, other members can present different views and justifications (Piaget's, 1985, genuine argument). Similarly, when Jay shows a gap in understanding by asking a question, other members can respond with explanations and justifications (Piaget, 1985). As justifications support the validity of an idea, they might help create CCs.

Group problem solving actions that hinder micro-creativity

Research on politeness suggests that disagreeing politely might aid creation of CCs but disagreeing rudely might hinder creation of CCs. When arguments spill over from the cognitive arena into the social arena, groups members might protect their public self-images (*face*) rather than further the problem solving (Brown & Levinson, 1987). Status differences can further aggravate these face concerns.

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, 2000). Evaluations range from polite to rude: agreement, neutral, change of topic, and disagreements (Chiu, 2000). Consider Jay's utterance again, "five times six is eleven." If Kay agrees with Jay ("right"), she supports him, promotes his face, and enhances their social relationship, especially if they invest themselves in their ideas (Brown & Levinson, 1987). Thus, members often agree with one another and repeat shared information to create common ground and solidarity (Chiu, 2000).

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 can be rude if the previous speaker (Jay) expects a response. For example, if Jay asks, "three times four is seven?" and Kay says "I'm hungry," she either ignores him or does not listen to him, both of which are rude. Lastly, disagreements (e.g., “no, you’re wrong”) can threaten face by lowering public perception of the previous speaker's (Jay’s) competence (Brown & Levinson, 1987).

When a person disagrees (e.g., Kay), the target person (Jay) ideally tries to understand the criticism and use it productively to create a CC. However, the threat to Jay's face may ignite his impulse to retaliate emotionally (*face attack*, “no, you're wrong,” Chiu & Khoo, 2003). Thus, rude disagreements threaten face, escalate interpersonal conflict, and often hinder creation and recognition of CCs. In this hostile environment, group members might withhold CCs or correct evaluations rather than risk losing face (Chiu, 2000). In the worst case, a spiral of rude disagreements can end the collaboration (Chiu & Khoo, 2003).

To avoid threatening Jay’s face, Kay might go to the opposite extreme and publicly agree. By doing so, Kay enhances her social relationship with Jay at the expense of their problem solving. Such false agreements allow errors to persist and potential CCs to remain unspoken. Also, group members might avoid disagreeing with one another due to pressure to achieve premature consensus (Janis, 1989).

Avoiding the extremes of rude disagreement and false agreement, Kay can disagree politely to reduce the threat to Jay's face while maintaining problem solving integrity (Chiu & Khoo, 2003). Instead of "no, you're wrong," Kay can disagree politely, "If five is multiplied by six, we don't get seven." The polite disagreement both reduces blame and creates common ground. First, Kay uses the hypothetical "if," thereby distancing the idea from reality. Second, she does not assign blame (no “you”). Third, Kay uses the passive voice, "is multiplied," to hide causal agency and responsibility. Lastly, she uses the passive circumstantial verb "get," thereby implicating agency in external conditions.

Kay’s polite disagreement creates common ground by repetition and shared positioning. By repeating Jay’s computation, "five is multiplied by six . . . eleven," Kay suggests that she shares his understanding. Also, Kay uses shared positioning, specifically the first person plural pronoun "we," to claim common cause with Jay.

Kay’s polite disagreement supports her relationship with Jay, so he is less likely to retaliate. Instead, Jay is more likely to try to understand Kay's criticism, recognize the flaw, and correct it with a CC (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 can support both the social relationship and the micro-creativity.

Status

Status differences can reduce CCs and distort evaluations of CCs through status struggles (Bales, 2001) or through the greater influence of high status members (Cohen, 1994). 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). During status struggles, intentional rude disagreements (*face attacks*; e.g., "everyone knows that five times six is thirty, not eleven") can enhance one's own face at the expense of a competitor's face (Chiu & Khoo, 2003). As noted earlier, rude disagreements can hinder creation of CCs.

After a status hierarchy has been established, status affects the expectations of individual group members (Cohen, 1994). In *expectation states* theory, status is linked to the expectation of competencies for the current activity (Cohen, 1994). 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, thereby distorting evaluations to agree excessively with high status people’s ideas and disagree excessively with low status people’s ideas. By doing so, members enact their expectations of high status members dominating the interaction and increase the ratio of flaws to correct ideas in the group’s shared knowledge base. High status people's influence can also increase over time. High status people speak early and often (Hackman & Johnson, 2000). As group members value and prefer supporting previously discussed or shared information rather than introducing unshared information, high status people's domination increases in severity over time (Stasser & Birchmeier, 2003).

Greater status differences can exacerbate status effects or raise the incentives for status struggles, both of which might reduce CCs. For group problem solving, the primary status characteristic is often past task achievement, but group members might also use diffuse status characteristics (such as social status) to make assumptions about one another’s competence.

In short, we have the following micro-creativity hypotheses.

H1. Groups with more diversity (race, gender) create more CCs.

H2. New ideas, whether correct or incorrect, help create CCs.

H3. Social metacognition (correct evaluations, disagreements, questions) facilitates CC creation.

H4. Rude disagreements and false agreements hinder creation of CCs while polite disagreements aid the creation of CCs.

H5. Greater status differences (achievement, social status) reduce CCs.

Method

Participants

Eighty students in an urban US high school participated in this study. The students attended four ninth grade algebra classes for seven months. On a state-wide exam, the school scored between the 40th to 50th percentiles in mathematics. These students had not received any group training and had not worked together.

All 80 students answered two questions regarding social status, "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." Then, these students were placed into groups of four. There was no same gender or same race groups. There were 40 girls and 40 boys; their races were 12 Asian, 27 Black, 28 Hispanic, and 13 White.

Procedure

The teacher presented the following problem to the students 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?"

The teacher used this problem to introduce them to a new unit on algebraic equations with multiple variables. Hence, the students had not yet learned, in class, any procedures for solving this problem. Furthermore, the problem involved complicated mathematics relationships, non-trivial combinations of multiple operations, and many solution methods. One solution method is equating the distance computations for each vehicle, cruise ship and helicopter (22 mph x [Time + 5 hours] = 90 mph x Time), to obtain 1.618 hours or 1 hour 37 minutes.

The students worked in groups for 30 minutes. They had pens, paper, and calculators available for their use. All students were videotaped, and the subsequent data were transcribed and coded.

Variables

This study included the following individual, group, and time period variables. Students identified themselves as *girls* or boys and as *Asian*, *Black*, *Latino* or White. *Mathematics grade* refers to students' mid-year algebra grades. Each student's *peer friendship* measure was the mean number of times his or her name appeared on classmates' answers to two social status questions. Group variables included the *group means* and *variances* of the above variables. Each group's final answer was coded as correct or wrong (1 or 0; *solution score*). Each group's problem solving session was also divided into *time periods* according to the DMA method described below. The number of time periods in a group ranged from one to five.

The transcripts of the videotapes were divided into sequences of words or actions by the same person (speaker turns). Blind to the study's hypotheses, two research assistants coded all the transcripts' speaker turns. Krippendorff's (2004) α was used to test for inter-coder reliability.

A speaker turn is a sequence of words by a group member bracketed by the words of other group member(s). A turn could also consist of only non-verbal actions (e.g., writing "3 x 40"). In contrast to flat classification schemes that only allow one or two codes for each speaker utterance (e.g., Bales, 2001), the research assistants coded each turn along the five dimensions suggested by these turn properties: evaluation of the previous action (EPA: *agree*, *ignore*, *rudely disagree*, politely disagree), knowledge content (KC: new idea [*contribution*], old idea [*repetition*], null problem content), validity (*correct*, *wrong*, null problem content), *justification*, and invitational form (IF: *command, question*, statement), (Chiu, 2000). A correct contribution (CC) is a new idea relative to the group members that is consistent with both the problem situation and algebra. With limited knowledge about the group members' experiences, a turn was coded as a contribution if it was not in the problem statement, not in the textbook, and not discussed earlier during the lesson.

Data Analysis

Statistical analysis of group processes at the speaker turn/utterance level is problematic as it involves time-series data and multiple groups. Time-series data collected from different groups often violate the independence assumption. Also, the effects of the explanatory variables on the outcome variable can change over time (non-stationarity) or differ across groups (group heterogeneity). This study uses a new method, *dynamic multi-level analysis* (DMA, Chiu & Khoo, 2005) to solve all of these difficulties: (a) identify distinct time periods, (b) test for group heterogeneity and time period stationarity, (c) multilevel analysis of the binary outcome variable, (d) test for residual serial correlation, and (e) identify direct and indirect effects.

Within a problem solving session, there might be fewer CCs at the beginning time period when people are trying to understand the problem than at later time periods when they are close to a solution. Hence, DMA identified watershed *breakpoints* that divided the time series data into time periods with significantly more or fewer CCs. For each group, I used the Bayesian information criterion (BIC) in a modified version of Maddala and Kim's (1998) method. The BIC measured whether a model has both goodness of fit and a parsimonious specification. In a regression of the outcome variable, CC, possible breakpoints serve as independent variables, and the BIC is computed for a simple time-series model (an auto-regressive order 1 model).

$$y_t = C + C_2 d_2 + C_3 d_3 + \ldots + C_p d_p + \beta y_{t-1} + \varepsilon_t \quad (1)$$

The parameter y_t indicates the value of the outcome variable y at speaker turn t. The parameter y_{t-1} indicates the value of the outcome variable in the previous turn, and β is its regression coefficient. Meanwhile, C is a constant, and ε_t is the residual at turn t. Each time period p has a dummy variable d_p and a regression coefficient C_p. Assuming zero to five breakpoints and calculating the BICs for all possible locations of those break points, the optimal model has the lowest BIC. Applying this method to each group yielded the number and locations of breakpoints (and hence time periods) for each group.

A multi-level Logit model can be divided into its multi-level part and its Logit part. Consider a 3-level model with an outcome variable, y_{ijk} (CC) at speaker turn i of time period j in group k and a Logit link function:

$$y_{ijk} = \beta_{000} + e_{ijk} + f_{0jk} + g_{00k} \quad (2)$$

$$\pi_{ijk} = p\,(y_{ijk} = 1) = F\,(\beta_0 + f_{0jk} + g_{00k}) = \frac{1}{1+e^{-(\beta_0 + f_{0jk} + g_{00k})}} \quad (3)$$

The level-2 variation parameter f_{0jk} is the deviation of time period j from the overall mean, and g_{00k} is the deviation of group k from the overall mean β_{000}. The probability (π_{ijk}) that an event (e.g. a CC) occurs at turn i of time period j in group k is a function of the expected value of the outcome variable and the Logit link function. As the level-1 variation, e_{ijk}, does not contribute to the fixed components and is a random variable only at level 1, it is constrained to 1 without loss of generality.

$$y_{ijk} = \pi_{ijk} + e_{ijk}[\pi_{ijk}(1-\pi_{ijk})]^{0.5} \quad (4)$$

The full analysis with all vectors of variables is shown in equation 5.

$$\pi_{ijk} = F(\beta_0 + \beta_{00s}\mathbf{S}_{00k} + \beta_{00t}\mathbf{T}_{00k} + \beta_{ujk}\mathbf{U}_{ijk} + \beta_{vjk}\mathbf{V}_{(i-1)jk} + \beta_{vjk}\mathbf{V}_{(i-2)jk} + \beta_{vjk}\mathbf{V}_{(i-3)jk} + \beta_{vjk}\mathbf{V}_{(i-4)jk} + f_{0jk} + g_{00k}) \quad (5)$$

First, a vector of s classroom identification control variables was added (**S**). As the likelihood ratio test for the significance of additional explanatory variables was not reliable for this estimation method, Wald tests were used (Goldstein, 1995). Non-significant variables were removed from the specification at each stage.

Then, t variables at the group level were added: correct group solution, mean of group members' mathematics grades, mean of group members' social statuses, variance of mathematics grades, variance of social statuses, racial variance, gender variance (**T**). The last two pairs of variables tested the status effects hypothesis (H5) and the diversity hypothesis (H1). Then, interactions among pairs of significant variables in **T** were added. Next, u current speaker variables was added: gender, race, mathematics grade, social status, correct evaluation, agree, politely disagree, rudely disagree, justify, question and command (**U**). Then, the speaker turn level regression coefficients were tested for significant differences (Goldstein, 1995).

I entered lag variables for the previous speakers (namely, a vector autoregression or VAR, Kennedy, 2004), first lag 1 (indicating the previous turn and denoted -1), then at lag 2 (denoted -2), then at lag 3, and so on until none of the variables in the last lag were significant (lag 4 in this case). First, I added v previous speaker variables at the speaker turn level: gender (–1), race (–1), mathematics grade (–1), social status (–1), correct evaluation (–1), agree (–1), politely disagree (–1), rudely disagree (–1), CC (–1), wrong contribution (–1), correct old idea (–1), justify (–1), question (–1), and command (–1) (**V**). These variables tested the new ideas, social metacognition, and rudeness hypotheses (H2, H3, and H4). Then, lags –2, –3, and –4 of the variables in **V** were added. The parameters were estimated with predictive quasi-likelihood methods (Goldstein, 1995).

Ljung-Box (1979) Q-statistics were used to test for serial correlation (up to order 4) in the residuals. If the residuals are not serially correlated, the parameter estimates are likely efficient, and the standard error estimates are likely unbiased (Kennedy, 2004). Otherwise, lagged outcome variables (lags of CC) can be added as explanatory variables, or the serial correlation can be modeled (Kennedy, 2004).

Based on the multilevel analysis results, the path analysis estimated the direct and indirect effects of the significant explanatory variables (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 total effects of each predictor were converted to odds ratios, reported as the percentage increase or decrease (+ X% or – X%) in the likelihood of a CC (Kennedy, 2004).

An alpha level of .05 was used for all statistical tests. The false discovery rate (FDR) was controlled via Benjamini, Krieger, and Yekutieli's (2006) two-stage linear step-up procedure, as their computer simulations showed that it was superior to 13 other methods. I also estimated the accuracy of the final model's micro-creativity predictions in each turn.

Results

As shown in an earlier analysis of this data, groups with a larger percentage of CCs had higher solution scores (Chiu & Khoo, in press). Thus, DMA was used CC creation. DMA identified watershed breakpoints that distinguished time periods of frequent CCs and infrequent CCs. Preliminary examination of the breakpoints suggest three broad types (on-task ↔ off-task transitions, insights, and critical errors). These time periods were incorporated into the explanatory models of CCs. Krippendorf's alpha for evaluations of previous actions, knowledge content, correctness, and invitational form were 0.93, 0.98, 0.99, and 0.91 respectively.

Supported Hypotheses
H2. Wrong, new ideas aid micro-creativity.
H3. Social metacognition (correct evaluations and disagreements –if polite) aids micro-creativity.
H4. Rude disagreements and false agreements hinder micro-creativity.
Polite disagreements aid micro-creativity.

Group properties did not affect the likelihood of a CC. CC creation was not linked to greater racial diversity, gender diversity, or status differences (showing no support for H1 or H5).

In contrast, recent speaker actions (constituting the local time context) affected the likelihood of CCs. Wrong, new ideas yielded more CCs (+9%; see Figure 1), but prior CCs did not yield more CCs, showing partial support for H2. After a wrong idea, group members were less likely to agree (–18%) and more likely to rudely disagree (+6%). Furthermore, they often detected and corrected errors to create a CC (e.g., "should be two hours, not five hours"). Surprisingly, a CC did not help create subsequent CCs. Thus, these groups had few chain reactions of CCs in this study.

When group members evaluated correctly or justified ideas, the likelihood of CCs increased, so the results partially supported H3. If any of the three previous speakers evaluated correctly (–1, –2, –3), the current speaker was more likely to create a CC (+2%, +5%, and +3%, respectively). Furthermore, a group member who evaluated correctly often helped subsequent speakers evaluate correctly, both in the next turn and in the following turn (+12% and +14%, respectively). Also, group members who evaluated correctly helped other group members justify their ideas (+3%), create fewer wrong ideas (–3%) and agree more often (+2%).

When group members justified ideas, the likelihood of a CC rose substantially in unsuccessful groups (+29%) and rose even higher in successful groups (+70%). This difference might stem from successful and unsuccessful group members' different sources for justifying their ideas. Successful group members often referred to the "RTD" (rate × time = distance formula) and other mathematics formulas to justify their ideas. In contrast, unsuccessful group members often referred to the teacher "because Ms. T said so," which might be less valid or less helpful to other group members. When a group member justified an idea, other group members were less likely to disagree rudely (–5%) and more likely to follow with justifications (+4%). Meanwhile, asking a question did not significantly affect the likelihood of a CC, and the effect of disagreeing depended on whether a group member disagreed politely or rudely.

Disagreements did not always increase CC creation (no support for this part of H3). Instead, polite disagreements increased CC creation (+14%), while rude disagreements reduced it (–4%). Controlling for correct evaluations, excessive agreement also reduced CC creation (–5%). These effects all supported H4. (Other predictors were not significant.)

Only two predictors showed different effects across time periods (agree and correct evaluation [–2]). Agreements yielded fewer CCs, with the effects varying across time periods from –3% to –21% (–5% overall). Correct evaluation (–2) generally yielded more CCs, with the effect varying across time periods from –0.3% to +9% (+5% overall). The varying effect sizes of agreement and correct evaluations across time periods suggested that their effects were moderated by unexamined variables that differed across these time contexts. Aside from justifications, agreements, and correct evaluations (–2), the effects of all other predictors did not differ significantly across time periods or across groups, and thus showed no evidence of contextual effects.

This model had an 83% accuracy rate for predicting whether a CC occurred in any given turn (y_{ijk}* vs. y_{ijk}). Furthermore, the Q-statistics showed no significant serial correlation of residuals in any of the twenty groups. So, the time-series model was likely appropriate.

Discussion

Diverse views and social metacognition might facilitate creativity while concerns over public self-image (face) or status might hinder it (Paulus & Brown, 2003). This study tested the group micro-creativity versions of these hypotheses using a statistical discourse analysis. While a group's local time context affected its micro-creativity, a priori individual and group characteristics did not, highlighting the importance of examining actual group member actions and interactions rather than inferring them from coarser group or individual measures. Specifically, the likelihood of a correct, new idea (correct contribution or CC) was higher after a (a) wrong, new idea, (b) correct evaluation, (c) justification or (d) a politely disagreement. In contrast, rude disagreement and excessive agreement were both linked to lower CC creation. CC creation was not linked to racial diversity, gender diversity, or status differences. The effects of justifications differed across groups, while those of agreements and correct evaluations differed across time periods.

Greater CC creation after wrong ideas, correct evaluations, and polite disagreements suggest that social metacognition often aids micro-creativity. After a wrong idea, group members agreed less often and rudely disagreed more often, suggesting that they detected the flaw in the wrong idea. Furthermore, they often corrected the flaw to yield a CC. Thus, wrong, new ideas might have served as kindling for micro-creativity via sparked ideas, jigsaw pieces, or creative misinterpretations (Chiu, 1997). In contrast, a CC did not raise subsequent CC creation in this study, suggesting that chain reactions of micro-creativity might not be common in similar situations. Unlike CCs, correct evaluations had long-term effects on CC creation. Correct evaluations increased subsequent correct evaluations, justifications, and CC creation over three turns. These results are consistent with the view that verifying correct ideas or identifying flaws helps create a shared, valid knowledge base for subsequent group micro-creativity.

Meanwhile, justifications reduced rude disagreements and yielded more justifications and CCs, suggesting that justifications aid rational discourse and micro-creativity. By encouraging polite discussions, justifications focus attention on the validity of an idea and encourage further justifications and CC creation.

Social metacognition did not always raise micro-creativity, likely because group members are concerned about face. While disagreements can identify obstacles for the group to address (Piaget, 1985), rude disagreements reduce CC creation, whereas polite disagreements increase CC creation. Furthermore, excessive agreement yielded fewer CCs, consistent with the view that group members agreed excessively to enhance their social relationships at the expense of their micro-creativity (Janis, 1989).

This study showed significant differences across groups and across time periods. Successful groups had justifications that were more effective in yielding CCs. Meanwhile, many CCs were clustered together as time period variation accounted for 79% of the total variance. Also, the effect sizes of agreement and correct evaluations varied across time periods.

If future studies validate these results across problem content, contexts, and group histories, this exploratory study might have the following theoretical, methodological, and practical implications. Theoretically, this study suggests that understanding group micro-creativity requires explicating the influence of both the time period and the local time context. Watershed breakpoints (on-task ↔ off-task transitions, insights, and critical errors) separated time periods of high micro-creativity from those of low micro-creativity. Furthermore, the local time context of recent group member actions (wrong ideas, correct evaluations, justifications, agreements, and rude disagreements) influenced the degree of micro-creativity. Methodologically, a new statistical discourse analysis (dynamic multi-level analysis, DMA, Chiu & Khoo, 2005) identified the watershed breakpoints and aided testing of the micro-creativity hypotheses. With this new statistical tool, scholars can test many categories of hypotheses about people's conversations (or more generally, interactions).

Practically, these results suggest that group members (especially group leaders) might improve their micro-creativity by encouraging one another to express their ideas, justify them, and evaluate them carefully and politely. To aid free expression of ideas (including wrong ones), group leaders can create a safe and supportive group culture. Within this supportive culture, group leaders can elicit justifications and slow contemplation to aid correct evaluations and reduce impulsive, rude disagreements. Through these theoretical, methodological, and practical contributions, this study might help group members improve their micro-creativity.

References

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

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

Brown, P, & Levinson, S. C. (1987). *Politeness.* New York: Cambridge University Press.

Chiu, M. M. (1997). *Building on diversity*. Los Angeles: UCLA. (ERIC No. ED 410 325).

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

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*.

Cohen, E. G. (1994). Restructuring the classroom. *Review of Educational Research, 64*, 1-35.

Goldstein, H. (1995). *Multilevel statistical models.* Sydney: Edward Arnold.

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.

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.
Ljung, G., & Box, G. (1979). On a measure of lack of fit in time series models. *Biometrika, 66,* 265-270.
Maddala, G. S., & Kim, I. M. (1998). *Unit roots, cointegration and structural change.* New York: Cambridge University Press.
McLeod, P. L., Lobel, S. A., & Cox, T. H., Jr. (1996). Ethnic diversity and creativity in small groups. *Small Group Research, 27*, 2, 248-264
Orlitzky, M., & Hirokawa, R. Y. (2001). To err is human, to correct for it divine. *Small Group Research*, *32*, 313-341.
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.
Piaget, J. (1985). *The equilibrium of cognitive structures.* (T. Brown & K. L. Thampy, Trans.). Chicago: University of Chicago Press.
Sawyer, K. (2003). *Group creativity*. Mahwah, NJ: Erlbaum.
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.
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.

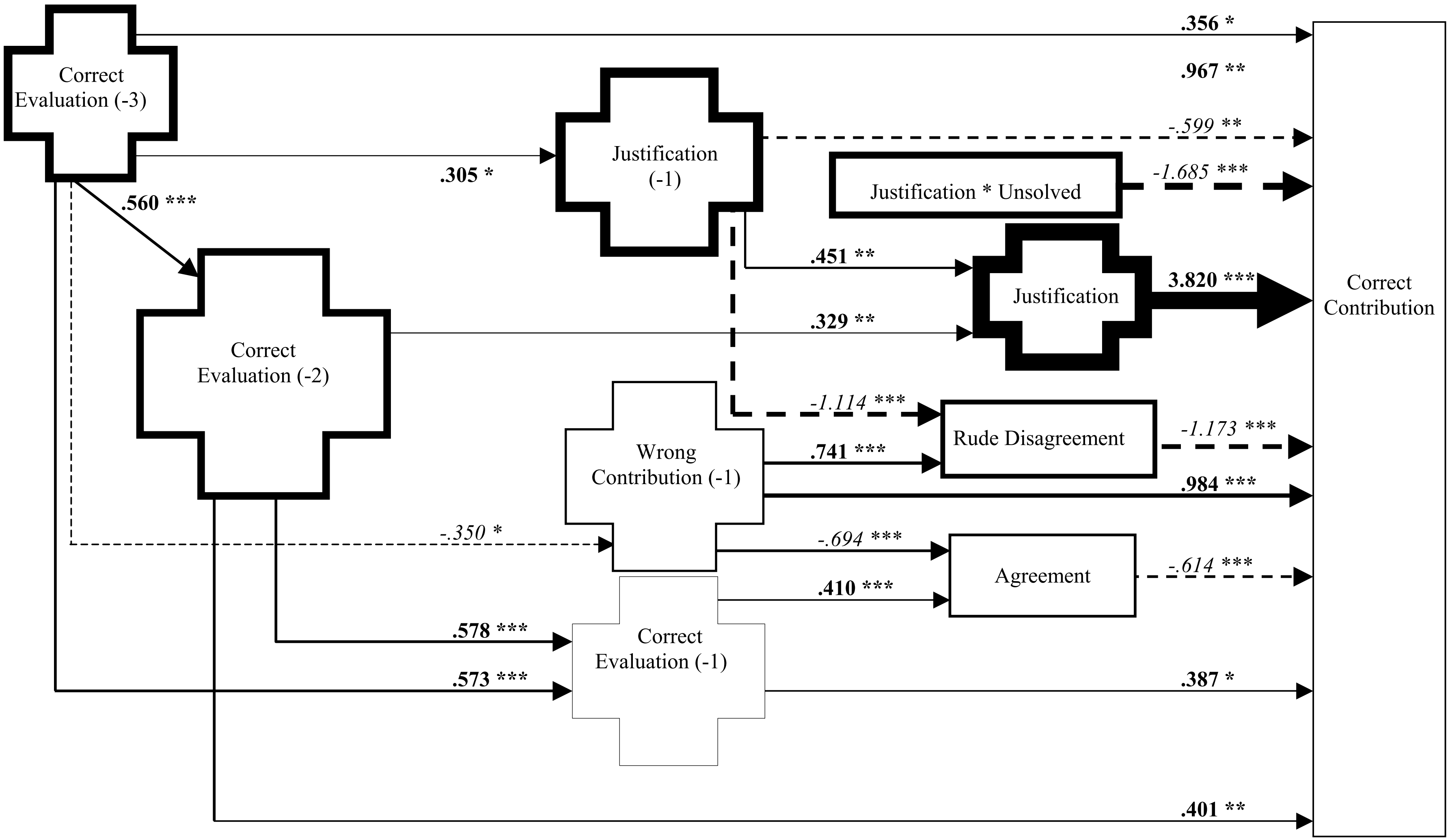

Figure 1. Significant predictors of likelihood of a correct contribution (correct, new idea; CC) in a path analysis with standardized parameters. Crosses (✚) indicate positive total effects, and rectangles (▫) indicate negative total effects. Solid arrows (→) indicate positive direct effects, and dashed arrows (-----) indicate negative, direct effects. Wider lines indicate larger effects.