Learning in Groups: Possible Advantages of Working in Threes

Jennifer Wiley (jwiley@uic.edu)
Department of Psychology, University of Illinois at Chicago
Chicago, IL 60607 USA

Lena Hildenbrand (lhilde3@uic.edu)
Department of Psychology, University of Illinois at Chicago
Chicago, IL 60607 USA

Thomas Ståhl (tstahl@uic.edu)
Department of Psychology, University of Illinois at Chicago
Chicago, IL 60607 USA

Abstract

This study investigated how interacting in different size discussion groups can improve students' understanding of research methods topics at the university level. Students completed a discussion activity in groups of twos, threes, and fours, and later were tested individually for their understanding of the target concepts. While the largest group size (fours) performed best on the worksheet that was completed in groups as part of the discussion activity, working in a group of three appeared to support better understanding of the target concepts for weaker students in the course. Several alternative explanations for the benefit from working in threes are considered.

Keywords: collaboration; learning; small group discussion activities; understanding; conceptual change.

Introduction

Despite the intuitive appeal of the idea that *two heads are better than one* and the popularity of educational activities based in small group interactions, few experimental studies have explored the effects of group size on both collaborative products and individual learning outcomes using direct comparisons of different group sizes. One popular reason why people believe groups are effective is the assumption that each group member brings to the table a slightly different set of task-relevant knowledge and skills (Canham, Mayer, & Wiley, 2012; Jarosz, Goldenberg, & Wiley, 2017; Maier, 1967; Wiedmann, Leach, Rummel, & Wiley, 2012). Through discussion, the knowledge and skills of each member can become available for all, and there is a higher likelihood that at least one individual will possess the knowledge needed to answer any given question (Hoffman, Harburg & Maier, 1962; Nemeth, 1986; Paulus & Yang, 2000; Stasser, Stewart, & Wittenbaum, 1995). From this perspective, one could predict a positive relation between group size and performance, with performance increasing with the size of a group.

From a constructivist learning perspective, the theoretical advantages of collaborative activities are based in mechanisms of cognitive development: where learning is prompted by discussion with others, or modeling of or for others (Piaget, 1932; Vygotsky, 1978). An aspirational goal for small group learning activities is that they might support students in reaching higher levels of reasoning or understanding than they could accomplish as individuals, and that learning in groups might help to reduce the gap in effectiveness between conventional instruction in which students learn on their own vs. learning from one-on-one tutoring (Bloom, 1984). The hope is that peers can serve some of the same functions as tutors (Brown & Palincsar, 1989; Webb, 1980; Webb, Troper & Fall, 1995) by increasing the amount of information or knowledge that is available (Canham, Mayer, & Wiley, 2012; Hoffman, Harburg & Maier, 1962; Maier, 1967; Nemeth, 1986; Paulus & Yang, 2000, Stasser, Stewart, & Wittenbaum, 1995), as well as by helping learners to plan, monitor and reflect on complex tasks or concepts (Coleman, 1998; Moreland & Levine, 1992; Roschelle & Teasley, 1995; Schoenfeld, 1989). Thus, from this perspective, one could predict that working in groups may especially benefit weaker students who are less likely to be able to complete instructional activities correctly on their own.

However, work on information processing in groups has shown that collaboration does not always lead to superior outcomes in idea generation tasks (Hill, 1982; Steiner, 1972; Taylor, Berry & Block, 1958), and these findings have also been shown in collaborative learning contexts (Cohen, 1994; Moreno, 2009; Salomon & Globerson, 1989). There are several tensions that can exist when students work in groups, and some disadvantages may increase as group size increases. One concern is the amount of effort that students contribute may depend on the size of the group. A pair of students is a group size that many teachers might assume could maximize the amount of effort from each individual, as larger group sizes might allow for more social loafing or free-riding (Gigone & Hastie, 1997; Kerr & Bruun, 1983). Further, more time may be spent on off-task topics in larger groups (Dugosh, Paulus, Roland & Yang, 2000). Interacting in groups also increases the number of information streams that need to be processed. In the end, individuals may be more burdened and enjoy less intact cognitive processing when working in larger groups (Hinsz, Tindale, & Vollrath 1997; Stroebe & Diehl, 1994).

Another tension exists between the number of perspectives that are available and the likelihood of evaluation apprehension which may prevent students from contributing their unique knowledge to a group (Camacho & Paulus, 1995). The likelihood of conflict may also increase with group size. Yet, while conflict may seem inherently negative, it can also provide the opportunity for meaningful discussion of ideas. Once there are more than two people in the group, this may increase the likelihood that a minority position exists (Moscovici, 1980; Nemeth, 1986) which in turn might prompt attempts to explain, justify, or evaluate different possible answers to questions. In turn, engaging in explanation processes may also led to more abstraction, re-representation or deeper conceptual understanding of the content, and the emergence of new ideas that no individual student held on their own (Schwartz, 1995; Webb, Troper & Fall, 1995). Chi's ICAP framework (2009; Chi & Wylie, 2014) suggests when students work together and engage in dialogue that this leads to learning in an *interactive* mode which is more effective than any other mode (passive, active or constructive).

From this perspective, a group of two may lead to a dynamic in which the members engage in little evaluation of each other's suggestions. For example, one study using a collaborative argumentation learning activity found disadvantages for students working in dyads (Wiley & Bailey, 2006). Instead of acting as critical evaluators of information and pushing the reasoning to a higher level, students learning in pairs generally engaged in passive acceptance of each other's proposals. This suggests that working in groups of three, more so than groups of two, might lead to deeper student understanding of concepts.

Consistent with these suggestions, some advantages of working in threes over twos have been found previously (Barron, 2003; Wiley & Jensen, 2006). Further, a pilot study on students enrolled in Introduction to Research Methods in Psychology examined outcomes when students completed recitation activities either as individuals or in groups with either two or three students. All other aspects of the course were held constant across conditions. Even though students were given the exact same worksheets during recitation, attended the same lectures, and engaged in the same scripted whole-class discussions, students who were assigned to work in triads for their recitation activities earned an average of 5 points higher in their exam grades than students who worked in either dyads or alone (partial eta squared = .05). Thus, the main result of this preliminary study supports the hypothesis that triads are a more effective group size than dyads for peer collaborative learning in this context.

It has not yet been tested whether students working in groups of four will show the same benefits as working in groups of three over students working in groups of two, and this is one main question motivating this study. To investigate this question, students completed a small group discussion activity as part of an Introduction to Research Methods course in Psychology. They were randomly assigned to work on a worksheet in groups of twos, threes, and fours. The topic of the activity was interpreting graphs showing main effects and interactions. The main dependent measure was individual student performance on exam questions on these target concepts which appeared on the next midterm. A second main question was whether weaker students might benefit more from different group sizes. Performance in the course was used to examine group size effects for the below-average students and the above-average students separately. The hypotheses tested in this study were: 1) if groups of three would lead to better student understanding than groups of two; 2) if groups of four would lead to similar advantages in student understanding, or even greater advantages than working in groups of three; and, 3) if group size effects would be more pronounced in learning outcomes of weaker students.

Method

Participants

University students (N = 487) who were enrolled in a 200-level course in Introduction to Research Methods in Psychology completed a small group discussion activity during one recitation section meeting of the course. Students received attendance points for participating. Students were randomly assigned to work on the activity in groups of two, three, or four. All students who did not participate in the small group discussion activity still took the midterm which included 5 exam questions on the target concepts. Students who did not complete the small group activity are included in the "individuals" category in Table 1. A power analysis based on the prior pilot study that found an advantage of working in groups of three over groups of two (Cohen's f = .23) suggested a target sample size of n = 75 students per condition in order to achieve an 80% probability of detecting a difference between group size conditions.

Table 1: Sample size and worksheet scores by group size.

Group Size	N	Worksheet Scores
Individuals	223	--
Groups of 2	91	78.4 (0.16)
Groups of 3	87	78.1 (0.14)
Groups of 4	86	84.1 (0.12)

Students were categorized as more-prepared or less-prepared based on their performance in the course prior to the activity. As shown in Table 2, we relied on random assignment to distribute stronger and weaker students across each of the three group sizes, $X^2(2, 264) = 3.74$, p = .15.

Table 2: Number of more-prepared and less-prepared students at each group size.

Group Size	More Prepared	Less Prepared
Individuals	102	121
Groups of 2	56	35
Groups of 3	42	45
Groups of 4	43	43

Materials

The main goal of this activity was to give students practice in how to identify patterns indicating the presence of main effects and interaction effects in line graphs showing fictitious results from 2x2 factorial designs. The activity took place as part of a discussion section associated with a lecture course in Introduction to Research Methods in Psychology. The discussion section meeting consisted of a brief introductory lecture about main effects and interactions, instructions for the group activity, filling out the group activity worksheet, and a post-activity wrap-up. Five exam questions tested the target concepts.

Background Prior to the discussion section meeting (the same week), two 50-minute lectures focused on the logic of factorial designs, main effects and interactions, and how to identify data patterns in line graphs and clustered bar charts indicative of each type of effect. The lecture included examples of finding main effects and interactions with the same kinds of line graphs that were the focus of the activity. All students, even those who did not attend the discussion section meeting, had access to the content from the lecture that was emphasized in the activity. At the start of the discussion section meeting, a teaching assistant (TA) provided a brief review of these lectures, focusing on how to identify main effects and interactions in a line graph. The same powerpoint slides from the lectures were used by all TAs. The review took about 10 minutes.

Worksheet Activity The worksheet consisted of eight line graphs, each from a 2x2 design (see Figure 1). For each line graph, the groups were instructed to specify whether the pattern of results was indicative of a main effect of variable A, a main effect of variable B, or an interaction between variables A and B. They were informed that more than one of these effects could be present in each graph. Students were given up to 30 minutes to complete the worksheet activity. There were 24 responses worth a total of 24 points.

Activity Wrap-up Upon completing the group activity, the TA collected the worksheets. The correct interpretations of all line graphs were then reviewed by the TA. All TAs used the same powerpoint slides for the wrap-up. The wrap-up took about 10 minutes, after which class ended.

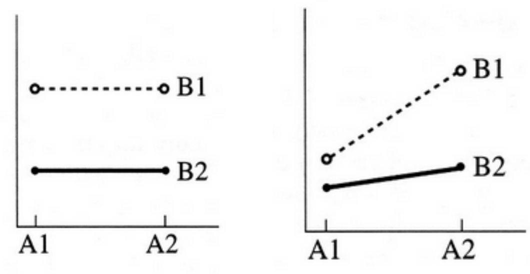

Figure 1: Example Line Graphs from Worksheet

Test Questions Understanding of the target concepts was tested with exam questions that presented students with line graphs showing the results of two studies, each with a 2x2 design, similar to those they had seen in lecture and the recitation activity. Students were provided with a scenario for each graph outlining the design of each study (Study 1: effects of two different drugs on psychotic symptoms among men and women; Study 2: effects of two different smoking cessation programs before and after treatment). Students were asked to describe the results in terms of main effects and interactions, and to draw conclusions about the efficacy of the drugs/smoking cessation programs. There were 5 open-ended questions worth a total of 5 points.

1. (Study 1) Based on the pattern of the means, what type of effect is most prominent in this case?

2. (Study 1) Based on the pattern of the means, what can you say about the effectiveness of the new drug as compared to the old drug?

3. (Study 2) Based on the pattern of means, what can you say about the effectiveness of each of the two programs?

4. (Study 2) Based on the pattern of means, there seems to be a problem with this study. What is the problem, and what caused it?

5. (Study 2) Come up with two alternative strategies through which this problem could have been avoided.

Procedure

The recitation section meeting began with the TA presenting background for the activity using slides from that week's lectures. The TA then randomly assigned students to work in groups of twos, threes, and fours. One worksheet was distributed to each group. The groups worked on the worksheet activity for up to 30 minutes. All worksheets were collected and the TA went over the correct answers. The students took a midterm exam that contained test questions on the target content. The exam took place 10 days after the discussion section activity. Scores on both the worksheets (24 possible points) and the exam questions (5 possible points) were converted to percentage correct.

Results

Worksheet Performance

Worksheet scores were analyzed using a 3x2 ANOVA looking at the effects of group size on performance for more-prepared and less-prepared students. As shown in Table 1, there was a significant difference in worksheet performance due to group size, $F(2, 258) = 5.26$, $p = .006$, $\eta^2_p = .039$. There was no difference in worksheet performance due to student preparedness, nor was there an interaction, *F*s <1.15.

Students working in groups of four received significantly higher scores on the worksheet than students working in groups of three or two. Groups of three and two did not differ. This result is consistent with a simple theoretical model of group size effects suggesting that larger groups may perform better because they maximize the chances that at least one member is able to provide a correct answer.

Exam Performance

Exam questions from the midterm following the activity tested students' understanding of how to interpret main effects and interactions from graphs. Performance on the target questions was analyzed using a 3x2 ANOVA looking at the effects of group size for more-prepared and less-prepared students. As shown in Figure 2, there was a significant effect of group size as performance on the target questions was better for students who participated in the discussion activity in larger groups, $F(2, 258) = 6.54$, $p = .002$, $\eta^2_p = .048$. Performance was also better for more-prepared students, $F(1, 258) = 51.01$, $p < .001$, $\eta^2_p = .165$. These two effects were additive and there was no significant interaction, $F(2, 258) = 1.18$, $p = .310$.

To follow-up the significant group size effect in the full sample, pairwise comparisons using LSD indicated that students who worked in groups of three ($p < .001$) and four ($p = .031$) performed better on the test questions than students who worked in groups of two. There was no evidence that working in groups of four led to additional advantages over working in groups of three. If anything, groups of four appeared to be somewhat less effective (albeit not significantly so, $p = .159$).

Separate Analyses for Less-prepared and More-prepared Students Although the interaction did not reach significance, due to *a priori* interest in which group size would be most effective for weaker students, separate analyses were performed for less- and more-prepared students. When effects were explored among just less-prepared students, there was a significant group size effect, $F(2, 120) = 4.23$, $p = .017$, $\eta^2_p = .066$. Working in groups of three led to significantly better exam performance than working in groups of two ($p = .005$). Performance for students who worked in groups of four did not significantly differ from groups of two ($p = .249$) or three ($p = .076$). In contrast, exam performance did not significantly differ due to group size among more-prepared students, $F(2, 138) = 2.94$, $p = .056$, $\eta^2_p = .041$. Thus, the advantage of working in groups appears to be more robust for less-prepared students.

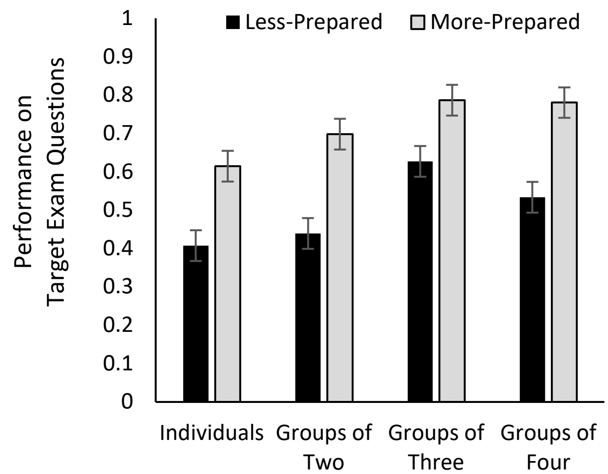

Figure 2: Average Exam Performance by Discussion Group Size and Student Preparedness (Error bars represent standard error).

Comparisons to Individuals An additional analysis explored effects of group size in an ANOVA that included the students who did not come to the group activity (individuals, not randomly assigned). Again, there was a significant effect for group size (now with four levels), $F(3, 479) = 16.36$, $p < .001$, $\eta^2_p = .093$, with advantages for working in groups over individuals at each group size (twos $p = .044$, threes $p < .001$, fours $p < .001$). However, a separate analysis on just less-prepared students found that individuals who did not come to the group activity performed similarly to students who attended the discussion activity and interacted in groups of two ($p = .452$), but less well than students who worked in groups of three ($p < .001$) or four ($p < .019$). More-prepared students showed advantages for working in groups over individuals at each group size (twos $p = .020$, threes $p < .001$, fours $p < .001$).

Discussion

Previous work has shown advantages in student understanding from working in small groups of threes compared to groups of twos (Barron, 2003; Wiley & Jensen, 2006). The main question motivating this study was extending this result to test whether students working in groups of four would show the same or greater benefits than working in groups of three. Another goal was to test whether discussion group size might be especially important for the learning outcomes of less-prepared students. The specific hypotheses tested in this study were: 1) if groups of three would lead to better student understanding than groups of two; 2) if groups of four would lead to similar advantages in student understanding, or even greater advantages than working in groups of three; and, 3) if group size effects would be more pronounced in learning outcomes of weaker students.

The first analysis considered performance on the group worksheets. The main result was that students working in groups of four received higher worksheet scores. Scores for groups of three and two did not differ. This result is consistent with a theoretical model suggesting that larger groups may perform better because they maximize the chances that at least one member is able to provide a correct answer.

In contrast, analyses performed on the target exam questions which provided a measure of how well each student understood the target concepts showed a different pattern of group size effects. Instead of a group size of four being best, students working in both groups of three and four showed evidence of better understanding than students who worked in groups of two. Further, weaker students who worked in groups of two did not show an advantage over the individuals, and only showed an advantage when working in groups of three. These results replicate prior work showing a benefit of working in threes over working in twos (Barron, 2003; Wiley & Jensen, 2006). They also extend this result in two ways.

First, they show that working in pairs may not actually result in any benefit to students (Moreland, 2010; Wiley & Bailey, 2006). Here, weaker students who completed the discussion activity working with just one other student did not perform significantly better than students who did not complete the activity at all.

Second, there was no evidence of the greatest benefit in understanding was associated with the largest group size. If anything, there was a trend for groups of four to be less effective at supporting student understanding. These results are most consistent with theoretical perspectives that emphasize both the advantages and disadvantages of larger group sizes. It may be that a group of three provides a sweet spot that maximizes the possible advantages while minimizing the disadvantages of working with others. It is also conceivable that it is not just having more than two members in a group, but also having an odd number of members that may contribute to the effectiveness of working groups of three. Working with an odd number of partners may increase the chances that there is a minority opinion. As previously discussed, this may help to prompt the kinds of meaningful interactions that can promote deeper conceptual understanding and construction of new ideas, as students attempt to explain, justify, or evaluate different proposed answers to questions (Chi, 2009; Jarosz, Goldenberg, & Wiley, 2016; Webb, Troper & Fall, 1995).

These results are of course limited in a number of ways. They were collected in an authentic course context, using a single discussion activity, and without attempting to control the membership of the groups or the way that students interacted with each other. This means there are a number of factors that could have influenced the results, including the specific topic, the background that students were given, and the design of the discussion activity. The results need to be explored with different topics, materials, and activity types to test if this pattern of results will generalize beyond this specific situation. For future work it will also be of interest to explore the composition of the groups in terms of the skills and knowledge of each of the members, and the mix of less-prepared and more-prepared members (Canham, Mayer, & Wiley, 2012; Jarosz, Goldenberg, & Wiley, 2016; Webb, 1980; Wiedmann, Leach, Rummel, & Wiley, 2012; Wiley & Jolly, 2003). Additional insights may be obtained by analyzing discourse patterns among students working in different group sizes. Understanding the group contexts and interactions that make some groups more effective than others could help develop tools for more tailored instruction and discussion support, ensuring that all students benefit.

In summary, assigning students to work in groups can enhance learning, particularly for less-prepared students. However, group size plays a crucial role -- only students in groups of three or four showed learning benefits, whereas weaker students who worked in pairs performed no better than individuals who did not participate in the group activity at all. These findings highlight the importance of structuring collaborative learning experiences thoughtfully, optimizing group size, and fostering productive discussions.

Acknowledgments

The authors thank Hannah Garcia, Taylor Strickland Miller, and Dhruti Yellanki for assistance coding the open-ended responses. This material is based upon work supported by the National Science Foundation under Grant No. 2120658. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

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

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, *13*(6), 4-16.

Brown, A. L., & Palincsar, A. S. (1989). Guided, cooperative learning and individual knowledge acquisition. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 393-451). Lawrence Erlbaum Associates.

Camacho, L. M., & Paulus, P. B. (1995). The role of social anxiousness in group brainstorming. *Journal of Personality and Social Psychology*, *68*(6), 1071-1080.

Canham, M. S., Wiley, J., & Mayer, R. E. (2012). When diversity in training improves dyadic problem solving. *Applied Cognitive Psychology*, *26*(3), 421-430.

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*(1), 73-105.

Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist, 49*(4), 219-243.

Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research*, *64*(1), 1-35.

Coleman, E. B. (1998). Using explanatory knowledge during collaborative problem solving in science. *Journal of the Learning Sciences*, *7*, 387-427.

Dugosh, K. L., Paulus, P. B., Roland, E. J., & Yang, H. C. (2000). Cognitive stimulation in brainstorming. *Journal of Personality and Social Psychology*, *79*(5), 722-735.

Gigone, D., & Hastie, R. (1997). Proper analysis of the accuracy of group judgments. *Psychological Bulletin, 121*(1), 149-167.

Hill, G. W. (1982). Group versus individual performance: Are *N* + 1 heads better than one? *Psychological Bulletin*, *91*(3), 517-539.

Hinsz, V. B., Tindale, R. S., & Vollrath, D. A. (1997). The emerging conceptualization of groups as information processors. *Psychological Bulletin*, *121*(1), 43-64.

Hoffman, L. R., Harburg, E., & Maier, N. R. F. (1962). Differences and disagreement as factors in creative group problem solving. *The Journal of Abnormal and Social Psychology, 64*(3), 206-214.

Jarosz, A. F., Goldenberg, O., & Wiley, J. (2017). Learning by invention: Small group discussion activities that support learning in statistics. *Discourse Processes*, *54*(4), 285-302.

Kerr, N. L., & Bruun, S. E. (1983). Dispensability of member effort and group motivation losses: Free-rider effects. *Journal of Personality and Social Psychology, 44*(1), 78-94.

Maier, N. R. (1967). Assets and liabilities in group problem solving: The need for an integrative function. *Psychological Review, 74*(4), 239-249.

Moreno, R. (2009). Constructing knowledge with an agent-based instructional program: A comparison of cooperative and individual meaning making. *Learning and Instruction*, *19*(5), 433-444.

Moreland, R. L. (2010). Are dyads really groups? *Small Group Research, 41*(2), 251-267.

Moreland, R. L., & Levine, J. M. (1992). The composition of small groups. *Advances in Group Processes*, *9*, 237-280.

Moscovici, S. (1980). Toward a theory of conversion behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 13, pp. 209-239). Academic Press.

Nemeth, C. J. (1986). Differential contributions of majority and minority influence. *Psychological Review, 93*(1), 23-32.

Paulus, P. B., & Yang, H. C. (2000). Idea generation in groups: A basis for creativity in organizations. *Organizational Behavior and Human Decision Processes*, *82*(1), 76-87.

Piaget, J. (1932). *The language and thought of the child*. Harcourt Brace Jovanovich.

Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In C. E. O'Malley (Ed.), *Computer supported collaborative learning* (pp. 69-97). Springer.

Salomon, G., & Globerson, T. (1989). When teams do not function the way they ought to. *International Journal of Educational Research*, *13*(1), 89-99.

Schwartz, D. L. (1995). The emergence of abstract representations in dyad problem solving. *The Journal of the Learning Sciences*, *4*(3), 321-354.

Schoenfeld, A. H. (1989). Ideas in the air: Speculations on small group learning, environmental and cultural influences on cognition, and epistemology. *International Journal of Educational Research*, *13*(1), 71-88.

Stasser, G., Stewart, D. D., & Wittenbaum, G. M. (1995). Expert roles and information exchange during discussion: The importance of knowing who knows what. *Journal of Experimental Social Psychology*, *31*(3), 244-265.

Steiner, I. D. (1972). *Group processes and productivity*. New York: Academic Press.

Stroebe, W., & Diehl, M. (1994). Why groups are less effective than their members: On productivity losses in idea-generating groups. *European Review of Social Psychology*, *5*(1), 271-303.

Taylor, D. W., Berry, P. C., & Block, C. H. (1958). Does group participation when using brainstorming facilitate or inhibit creative thinking? *Administrative Science Quarterly*, *3*(1), 23-47.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Webb, N. M. (1980). An analysis of group interaction and mathematical errors in heterogeneous ability groups. *British Journal of Educational Psychology*, *50*(3), 266-276.

Webb, N. M., Troper, J. D., & Fall, R. (1995). Constructive activity and learning in collaborative small groups. *Journal of Educational Psychology, 87*(3), 406-423.

Wiedmann, M., Leach, R. C., Rummel, N., & Wiley, J. (2012). Does group composition affect learning by invention? *Instructional Science*, *40*, 711-730.

Wiley, J., & Bailey, J. (2006). Effects of collaboration and argumentation on learning from web pages. In A. M. O'Donnell, C. E. Hmelo-Silver, & G. Erkens (Eds.), *Collaborative learning, reasoning, and technology* (pp. 297-321). Lawrence Erlbaum Associates.

Wiley, J., & Jensen, M. S. (2006). When three heads are better than two. In *Proceedings of the 28th Annual Meeting of the Cognitive Science Society* (pp. 2351-2356).

Wiley, J., & Jolly, C. (2013). When two heads are better than one expert. In *Proceedings of the 25th Annual Meeting of the Cognitive Science Society* (pp. 1242-1246).