American Educational Research Journal
Month XXXX, Vol. XX, No. X, pp. 1–30
DOI: 10.3102/0002831215618663

Improving Children's Competence as Decision Makers: Contrasting Effects of Collaborative Interaction and Direct Instruction

Xin Zhang
Richard C. Anderson
Joshua Morris
Center for the Study of Reading, University of Illinois at Urbana-Champaign
Brian Miller
Towson University
Kim Thi Nguyen-Jahiel
Center for the Study of Reading, University of Illinois at Urbana-Champaign
Tzu-Jung Lin
Ohio State University
Jie Zhang
Western Kentucky University
May Jadallah
Illinois State University
Theresa Scott
Jingjing Sun
Beata Latawiec
Shufeng Ma
Center for the Study of Reading, University of Illinois at Urbana-Champaign
Kay Grabow
Thomas Paine Elementary School, Urbana, Illinois
Judy Yu-Li Hsu
Center for the Study of Reading, University of Illinois at Urbana-Champaign

This research examined the influence of contrasting instructional approaches on children's decision-making competence. A total of 764 fifth graders, mostly African Americans and Hispanic Americans, from 36

XIN ZHANG is a doctoral student in psychology at the University of Illinois at Urbana-Champaign, 603 E. Daniel St., Champaign, IL, 61820; e-mail: xzhang85@illinois.edu. *Xin Zhang carries two lines of research. One line of research focuses on the role of collaborative learning and argumentation in promoting children's higher-order reasoning. The other line of research addresses cultural influences on children's social development during early adolescence.*

RICHARD C. ANDERSON is University Scholar and Professor Emeritus at the University of Illinois at Urbana-Champaign, Center for the Study of Reading. Anderson has two lines of research: comparative analysis of learning to read Chinese and English and the role of classroom discussion in children's cognitive and social development.

JOSHUA MORRIS is a PhD student in educational psychology at the University of Illinois at Urbana-Champaign. His research is focused on collaborative learning and how argumentative discourse is related to higher-order thinking.

BRIAN MILLER is an assistant professor in elementary education at Towson University in Maryland. His research includes studying how emotions affect science learning, the development of argumentative skills, and the use of diagrams in learning science.

KIM THI NGUYEN-JAHIEL is the lab manager and coordinator of the collaborative reasoning (CR) group.

TZU-JUNG LIN is an assistant professor of educational psychology at the Ohio State University. Her research centers on the influence of dialogic inquiry on children's and adolescents' cognitive and social development and the roles of teachers and peers in collaborative learning.

JIE ZHANG is an assistant professor of educational research at Western Kentucky University. Her research focuses on cognitive and psycholinguistic processes involved in language and literacy development and the role of classroom discussion in children's language, literacy, and reasoning development, especially for English language learners.

MAY JADALLAH is an associate professor in the elementary education program at Illinois State University. Her research examines the impact of collaborative discussion on children's cognitive processes and teacher's scaffolding strategies. Jadallah is also exploring the use of digital mapping technologies in upper elementary grades.

THERESA SCOTT is a school teacher in the Champaign, Illinois, school district. For 20 years, she was a teacher and trainer in a private school in Pennsylvania where she taught collaborative reasoning to her classes.

JINGJING SUN is an assistant professor of educational psychology at the University of Montana. Her research interests lie in the interactions of children's social, emotional, and cognitive development in the context of collaborative learning.

BEATA LATAWIEC is a visiting assistant professor in educational and school psychology at Wichita State University. Her research focuses on language and cognition, psychology of reading, language, and communication.

SHUFENG MA is a PhD student in educational psychology at the University of Illinois at Urbana-Champaign. Her research focuses on the effect of collaborative group work on children's oral narrative skills, productive use of academic vocabulary, and multilink causal reasoning.

KAY GRABOW recently retired after teaching for 43 years at Thomas Paine Elementary School in Urbana, Illinois. She taught fourth-grade students every year of her career and used collaborative reasoning strategies in the classroom for over 10 years.

JUDY YU-LI HSU is a doctoral student in educational psychology at the University of Illinois at Urbana-Champaign. Her research passion lies in supporting second-language learners' oral language development through peer collaboration.

classrooms in eight public schools serving children from low-income families completed a six-week unit on wolf management, using either direct instruction or collaborative groups, or were waited-listed controls. Analysis of children's essays on a topic unrelated to wolves revealed that students who participated in collaborative groups but not students who received direct instruction acquired decision-making strategies and transferred them to the novel task. Students in collaborative group work classrooms wrote essays that were significantly better than essays of students from direct instruction classrooms in each of the three aspects of decision making that were evaluated—considering more than one side of a dilemma, comprehensiveness of reasons, and weighing the importance of reasons. In contrast, direct instruction students performed no better than uninstructed control students.

Keywords: decision making, collaborative reasoning, direct instruction, argumentation

> In training a child to activity of thought, above all things we must beware of what I will call "inert ideas"—that is to say, ideas that are merely received into the mind without being utilized, or tested, or thrown into fresh combinations.
>
> Alfred North Whitehead (1929/1957, p. 193)

Decision making has been studied in a variety of disciplines, including economics and psychology (for a review, see Weber & Johnson, 2009), politics and public policy (e.g., Dye, 1992; Howlett, 2009), and health and medicine (e.g., Chapman & Sonnenberg, 2003; Hibbard, Slovic, & Jewett, 1997). Because of its universal importance in human affairs, as early as the 1960s, Engle (1960) declared that decision making should be the heart of social studies instruction for children and adolescents. She explained that social studies should go beyond simple facts to inform students of the reasoning required to reach decisions on public and private matters. Newman and Oliver (1970) maintained that when teaching social studies, teachers should foster students' capacity for justifying or challenging decisions. However, Jonassen (2012) recently concluded that intervention programs for improving children's decision-making competence are far from satisfactory.

The current study represents a renewed effort to improve children's decision making. We evaluated a six-week intervention called the Wolf Reintroduction and Management Unit, during which students were responsible for making a decision about whether a community should be allowed to kill a pack of wolves that alarmed many of its citizens. The unit integrates science, social studies, and the English language arts, incorporating instruction in the ecosystem, economy, and public policy. The unit was taught using one of two instructional approaches. One approach was collaborative

group work. This approach was compared to teacher-led whole-class direct instruction and regular instruction in control classrooms. The key goal of the current research was to determine whether children could transfer decision-making skills acquired from the wolf unit to a novel task far removed from issues of wolf reintroduction and management. The overarching hypothesis was that collaborative group work would lead to greater transfer than direct instruction. As we will detail later, the basis for this hypothesis is that collaborative group work positions students as active decision makers, whereas direct instruction places them in a passive role following the reasoning of their teacher.

Research on Decision Making

Decision making has been defined as the process of making choices among competing courses of action (Beyth-Marom, Fischhoff, Quadrel, & Furby, 1991; Raiffa, 1968; von Winterfeldt & Edwards, 1986). Decision making is higher-order thinking in that it "challenges the student to interpret, analyze, or manipulate information, because a question to be answered or a problem to be solved cannot be resolved through the routine application of previously learned knowledge" (Newman, 1990, p. 44).

Models of decision making generally are of two types, normative and descriptive (for a review, see Regenwetter et al., 2009). Normative models deal with logically consistent decision procedures. The purpose is to specify how a rational, informed person should identify the optimal option via an appropriate procedure, like probability and utility analysis. For example, if a choice has to be made between: (a) a sure gain of $240 or (b) a 25% chance to gain $1,000 and 75% chance to gain nothing (Tversky & Kahneman, 1981), assuming people are strictly rational, they would select the option that maximizes the expected value or utility. Utility is determined by a simple formula: the probability of each option times the value of each option. Here the utility of (b) is 250 (25% $\times$ 1,000 + 75% $\times$ 0 = 250), which is higher than that of (a); therefore, a rational person would choose option (b) under normative models of decision making.

The goal of descriptive models is to represent the actual process of how individuals make decisions. Descriptive models start with the assumption that individuals have biases that affect decision making, that they often employ rules of thumb, and that they are influenced by their emotions and personal experience (for a review, see Weber & Johnson, 2009). In fact, a substantial body of empirical research has documented that people's behavior often deviates from the performance considered normative in reasoning and decision-making tasks (for a review, see Stanovich, West, & Toplak, 2012). For instance, individuals may assess probabilities incorrectly, violate the axioms of utility theory, and ignore alternative arguments when evaluating a decision. In the previous example, Tversky and Kahneman

(1981) found that when considering options for gain, most people were risk averse and opted for the sure thing over the chance for getting a larger gain. In contrast, when similar choices were framed in a negative way—a sure small loss versus the chance of a larger loss—people became risk takers; they preferred the choice that might entail a larger loss over a smaller but sure loss. Some descriptive models acknowledge a rational component in individuals' decision making (for a review, see Gigerenzer & Gaissmaier, 2011). These models posit, for example, that people use fast-and-frugal heuristics, which may bring positive outcomes if used in an adaptive manner.

Decision making often has been assumed to be a competence like language that develops naturally through maturation and socialization (Beyth-Marom et al., 1991). Some studies suggest that children and adolescents are less skilled than adults in several aspects of decision making, including advice seeking, weighing pros and cons of options, and adaptive goal setting (e.g., Byrnes, 2002). Halpern-Felsher and Cauffman (2001) found that adults outperformed children in making decisions about several types of dilemmas. Adults were more likely to spontaneously consider the risks, benefits, and long-term consequences associated with different decisions. Klaczynski (2001) compared the decision making of early and middle adolescents. Middle adolescents performed closer to the normative ideal than early adolescents, which suggests that the use of analytic processes generally increases with age. Findings of studies comparing people of different ages are not completely consistent, however. In some studies, adults were shown to be equally biased or even more biased than children and adolescents (e.g., Klaczynski, Gordon, & Fauth, 1997). Though there seems not to be a final conclusion yet, existing evidence suggests that, as compared to adults, children and adolescents lack skills and dispositions required for thoughtful decision-making.

Instructional Interventions to Improve Decision-Making Skills

Instructional interventions have been created with the goal of improving children and adolescents' decision making, either through direct instruction or collaborative discussion. The interventions have been designed to teach a process of good decision making through three major approaches: (a) a pure decision-making course, (b) an issue-based socio-science curriculum with explicit decision-making procedures outlined, and (c) an issue-based socio-science curriculum with embedded decision-making elements.

Pure decision-making courses impart knowledge about probability and basic principles in decision making through direct teaching. An example is the GOFER course (Mann, Harmoni, Power, Beswick, & Ormond, 1988), which presents a five-step decision-making process: (a) goals clarification, (b) options generation, (c) fact finding, (d) consideration of effects, and (e) review and implementation. For socio-science curricula in which

improved decision making is a goal, similar decision-making procedures have been taught. For example, Ratcliffe (1997) expanded the five-step process into six steps: (a) identifying possible options, (b) identifying suitable criteria for comparing different options, (c) clarifying information known about possible options, (d) evaluating the advantages and disadvantages of each alternative against the identified criteria, (e) choosing an option based on the analysis undertaken, and (f) evaluating the decision-making process undertaken. She incorporated the six-step decision-making procedure into an energy conservation unit. A similar six-step scheme was used in the Decision Skills Curriculum (Spitzhoff, Ramirez, & Wills, 1982) with the purpose of increasing students' resistance to internal and external pressures to engage in destructive behavior.

Whether students could recognize or repeat the normative procedures that have been taught has been the major criterion for evaluating the effectiveness of these decision-making courses. For example, the Decision Knowledge Questionnaire used to evaluate the GOFER course contained multiple choice and open-ended questions like: What makes someone a really good decision maker? What is the difference between a simple decision and a thinking decision? In asking students to repeat abstract decision-making principles, the researchers made the dubious assumption that heightened awareness of normative decision-making procedures was equivalent to a thoughtful decision-making process. Moreover, although students who received training had better memory for decision-making principles in the short term (Mann et al., 1988), the positive effects soon faded away, as Ratcliffe (1997) reported that two months after the intervention, none of the participants recollected the detail of the decision-making principles. In addition to verbalizing rules, students' self-evaluations have been used to show benefits from courses (e.g., The decisions I make turn out well. When faced with a decision, I go along with what others suggest.). It is encouraging to see improved self-efficacy, but the measures fall short of providing evidence that students can actually engage in thoughtful decision making.

An embedded issue-based socio-scientific decision-making curriculum does not teach decision-making principles in a didactic manner; instead, it utilizes an integrated, issue-based approach to guide students toward informed decision making, with the aim of promoting their conceptual understanding as well as their ability to engage in rational argumentation. For example, Lee (2007) used the Hong Kong government's ban on smoking as the topic to engage students in argumentation and decision making. Teachers attempted to develop students' basic conceptual understanding about smoking in multiple ways, such as introducing information about the structure and function of the human respiratory system, doing experiments, and providing evidence showing how smoking affects health. Students were encouraged to explore both sides and evaluate their views

in a rational evidence-based manner. Besides smoking, other socio-scientific topics, such as animal transgenesis (Simonneaux, 2001) and bat conservation (Lee & Grace, 2010), have been employed in programs to improve decision making. Evaluations of these programs have typically had a pre- and posttest design where students were either interviewed or asked to write about their decisions; however, the topic was the same all the way through. In other words, there was not a novel transfer task, which makes it impossible to tell whether the participants had acquired generalizable decision-making skills that they are able to apply to new issues.

Collaborative Interaction and the Development of Higher-Order Thinking

Argumentative thinking inevitably accompanies deliberative decision making, and collaborative discussion has proved to be a successful method for promoting argumentative thinking (Kuhn, 1992; Reznitskaya, Anderson, & Kuo, 2007). Thus, we are among those who see collaborative interaction as a promising approach to promote decision-making skills (Braund, Lubeen, Scholtz, Sadeck, & Hodges, 2007; Lee, 2007; Simon & Maloney, 2007).

The expectation that collaborative interaction will foster children's decision-making skills is based on the premise that it intertwines social and cognitive processes that promote higher-order thinking. According to Piaget's (1932) socio-cognitive theory, cognitive development is initiated by disequilibrium between a child's current understanding and unsettling new information. During collaborative interaction, a child's existing understanding is often challenged by others, which creates a state of disequilibrium. In order to restore the equilibrium, the child needs to develop a new and integrated perspective to reconcile the conflict. In other words, cognitive conflict arising from social interaction brings children's thinking to a higher level of sophistication, thus facilitating cognitive growth.

Vygotskian socio-culture theory emphasizes the role of dialogical experience in intellectual growth. According to Vygotsky (1962), children's intellectual development from the actual level to the potential level relies largely on dialogical interaction with more expert partners. An expert partner can either be a knowledgeable adult or a competent peer. From the Vygotskian perspective, active exchange of ideas through verbal communication is critical in child's cognitive growth because language serves vital functions as a cognitive tool for the child to process information, as a social tool for sharing knowledge among children, and as a pedagogic tool to provide intellectual guidance (Mercer, 1999).

Over the past few decades, an accumulating body of research has provided empirical support for the benefits of collaborative interaction in promoting children's higher-order thinking in a variety of domains involving different age groups, including concept learning (e.g., Asterhan &

Schwarz, 2007, 2009), problem solving (e.g., Fawcett & Garton, 2005), skills of argument (e.g., Dong, Anderson, Kim, & Li, 2008; Dong, Anderson, Lin, & Wu, 2009; Kim, Anderson, Miller, Jeong, & Swim, 2011; Kuhn & Udell, 2003; Reznitskaya et al., 2001), and nonverbal reasoning (e.g., Mercer, 1996; Mercer, Dawes, Wegerif, & Sams, 2004; Wegerif, Mercer, & Dawes, 1999). A number of scholars have maintained that active exchange of ideas with others is the key mechanism in children's development of high-order thinking and reasoning (e.g., Rogoff, 1990; Teasley, 1995; Webb & Favier, 1999).

One approach to collaborative interaction is collaborative reasoning (CR), a peer-led free-flowing discussion forum intended to encourage authentic argumentation (Anderson, Chinn, Waggoner, & Nguyen, 1998). Students in heterogeneous small groups balancing school academic achievement, talkativeness, gender, and ethnicity are expected to find a resolution of an issue posed by a story that they have read (Anderson et al., 1998). They are supposed to manage their own discussions, and teachers intervene only when necessary. A series of studies have found that CR discussions foster children's acquisition of argument skills (e.g., Anderson et al., 1998, 2001; Lin et al., 2015). Students who have participated in as few as four or five CR discussions are able to compose essays about a topic they have not discussed with a more salient argument-counterargument-rebuttal structure than comparable control students (e.g., Dong et al., 2008, 2009; Kim et al., 2011; Reznitskaya et al., 2001, 2007; for a review, see Reznitskaya, 2009).

Rationale for the Current Research

Students spent six weeks studying the Wolf Reintroduction and Management Unit. The wolf unit can be classified as an embedded issue-based socio-scientific decision-making curriculum. The key message of the curriculum was to help young students become aware of how to make a reasonable and responsible decision. Through participating in the wolf unit, it was expected that students' decision-making skills would be fostered in at least three aspects: (a) disposition to consider more than one side of issues, (b) ability to generate different types of reasons on each side of issues, and (c) ability to weigh reasons on both sides. Children's decision-making skills were examined in a novel transfer task that had no surface features in common with wolf management, which overcomes a serious drawback of previous research on programs to teach decision making.

The wolf unit was taught in two formats: collaborative group work (CG) and direct instruction (DI). CG was a combination of collaborative reasoning discussions and other group activities. Students were broken into groups to discuss the "big question"— whether the community should be permitted to hire professional hunters to the kill wolves—and to discuss more specific questions related to the big question. DI entailed teacher-guided whole-class activities, whole-class question-and-answer sessions, and individual

seatwork. Wait-listed control classrooms continued regular instruction and received the wolf unit in the semester after the study.

Our expectation was that the collaborative groups would prove superior to direct instruction, as well as conventional instruction in control classrooms, on all three aspects of decision making. This expectation can be explained in terms of several differences in students' perspective and behavior during collaborative group work and direct instruction. A key difference is that, as realized in the Wolf Reintroduction and Management Unit, collaborative group work positioned the student as a decision maker personally responsible for reaching the best overall decision about the wolves. In contrast, in the direct instruction classroom, the student shared the decision-making responsibility with the teacher, and it was perhaps easy for the student to slip into the role of going along with what the teacher seemed to think was best. Because they individually and collectively shouldered the responsibility for decision making, students in collaborative groups may be considered to be active rather than passive learners. With a full sense of decision-making responsibility, we suppose, came the impetus to thoroughly explore all sides of issues.

In collaborative groups, each child is a decision maker invited to argue for his or her standpoint, justify, negotiate, and try to convince others (Dillenbourg, 1999), which provides a counterweight to myside bias (Wolfe & Britt, 2008) since participants are confronted with viewpoints different from their own. Possible benefits and harms of a decision, beyond the scope of any one individual's knowledge or imagination, may be brought to light by other participants. Hence, CG provides a context in which recurrent patterns of decision-making elements—considering the other side of the issue, appealing to different reasons supporting each side, and weighing reasons on each side—are likely to show up, which may bring children's understanding of decision making to a higher level of sophistication.

Although direct instruction has certain benefits (Kirschner, Sweller, & Clarke, 2006; Stein, Carnine, & Dixon, 1998), it may not foster students' higher-order thinking to the same extent as collaborative interaction. Teachers dominate talk in the conventional direct instruction classroom, and episodes of genuine dialogue between the teacher and students or among students are vanishingly small (Nystrand, Wu, Gamoran, Zeiser, & Long, 2003). Direct instruction is monological and hierarchical, limiting opportunities for students to engage in extended dialogue (Wells & Arauz, 2006), which according to Bakhtin (1981) is essential for the development of reasoning given its dialogical nature. During direct instruction, the teacher controls the topic, determines which students will have turns for speaking, and decides whether student contributions are satisfactory. When following the lead of the teacher, we hypothesize students are less likely to view themselves as real decision makers, thus attenuating engagement in the decision-making process.

Based on the idea that CG places students in the role of active decision makers, whereas DI positions them in a passive role conforming to the reasoning of their teacher, we hypothesized that CG students would show greater transfer of decision-making skills acquired from the wolf unit to a novel task than DI students. The novel transfer task had nothing to do with wolf management but addressed whether a boy should tell the teacher that a friend of his cheated in a model car competition.

Method

Participants

The participants were 764 fifth-grade students (410 girls, 354 boys), primarily African American (n = 338, 44.2%) or Hispanic American (n = 356, 46.6%), from 36 classrooms in eight public schools serving children from low-income families in two urban districts in the Midwest of the United States. From 79% to 99% of the students were registered for free or reduced-price lunch, depending on the school.

The 36 classrooms were divided into 12 triples of classrooms matched in terms of demographic composition and academic record. Classrooms within triples were randomly assigned to the three intervention conditions: collaborative groups (CG), direct instruction (DI), and wait-listed control (control). The 12 CG classrooms contained 254 children (143 girls, 111 boys), the 12 DI classrooms contained 256 (138 girls, 118 boys), while the 12 control classrooms contained 254 (129 girls, 125 boys).

Prior to the intervention, students took the Gates-MacGinitie reading comprehension test (MacGinitie, MacGinitie, Maria, & Dreyer, 2000). The normal curve equivalent–scaled scores on the reading comprehension test were below the national average (M = 38.37, SD = 16.24). Raw scores corrected for guessing, which improved reliability and validity, were employed in subsequent data analyses. Also prior to the intervention, students individually completed a speeded object naming task (Snodgrass & Vanderwart, 1980) to assess basic oral English proficiency. Students named common objects in two sets of pictures as quickly as they could; the score was number of pictures correctly named per minute. Two-level regression analyses with condition as a fixed effect and classrooms as a random effect showed no condition difference in reading comprehension, $F(2, 690) = 1.08$, $p = .34$, or basic English proficiency, $F(2, 690) = .06$, $p = .94$.

Wolf Reintroduction and Management Unit

In the wolf unit, students played the role of officials at the Wolf Management Agency. They were asked to make a decision about whether a community should be granted permission to hire professional hunters to kill a pack of wolves that concerned many of its citizens. The unit was

divided into three sections, each incorporating an important perspective on the complicated issue of wolves, to cultivate students' ability to discern different aspects of problems and understand interrelationships and trade-offs. The three sections were ecosystem, economy, and public policy. While killing the wolves may be favored by the majority of residents in the community (public policy), doing so would alter the food web (ecosystem), which would impact community businesses (economy).

Each section of the unit was explained in an information booklet and expanded in an activity booklet. Information booklets provided students with essential concepts. Unlike most readings for middle grade students, the booklets had an argument structure that contrasted opposing viewpoints. For example, the public policy information booklet compared the viewpoints of hunters who want a large elk population and nature lovers who want to hear howling wolves. The activity booklets contained problem-solving exercises that strengthened and expanded concepts that students were acquiring.

The wolf unit was designed to provide a balanced treatment, giving neither more nor less weight to granting permission to eradicate wolves or protecting them from being killed. In other words, the purpose of the curriculum was not to lead students to a predetermined best answer but rather to help young students become aware of how to make a responsible and reasoned decision that takes different perspectives into consideration. For example, wolves kill livestock such as cattle, sheep, and turkeys, which results in economic loss for ranchers and farmers. In order to protect livestock from being killed, ranchers may need to build fences and barns, which increases the cost of doing business. Therefore, allowing the wolf pack to be eradicated may seem to be the optimal choice for ranchers. However, the historical record indicates that livestock losses from wolf predation are small and, furthermore, the government compensates ranchers for losses they can document are due to wolves. Ranching is not the only business in the community. A campground, motel, and several restaurants benefit from tourism. Keeping wolves in the community may attract tourists from other parts of the country to see the wolves. On the other hand, family campers may stay away for fear that wolves will harm their children.

CG work was a combination of collaborative reasoning discussions and other group activities. Students were broken into groups to discuss the "big question"—whether the community should be permitted to hire professional hunters to kill the wolves. On a typical day during the unit, the task for a small group was to answer a specific question related to the big question, for example, "What effect would killing the wolves have on the elks?" Groups worked independently and spoke freely among themselves, with occasional assistance from the teacher. Each small group was assigned to become "experts" in one of the three domains of knowledge (ecosystem, economy, public policy). After four weeks of group work, each expert group

shared what they had learned in a poster presentation to the whole class. Then new discussion groups were created, with members from the three different expert groups, to reconsider the big question in a collaborative reasoning discussion.

DI entailed teacher-guided whole-class activities and individual seatwork. Students in DI condition sat facing toward the teacher. Students were supposed to raise their hands and wait for the teacher to select them before speaking. The teacher led students through all three domains of knowledge in the wolf unit. Students read the information booklets and completed the activity booklets individually at their seats. Students discussed the policy decision as a whole class.

Teachers who implemented the CG or DI interventions attended a two-day workshop to receive a detailed introduction to the Wolf Reintroduction and Management Unit and training in the method to which they were assigned. Teachers watched videos of the wolf unit as it had been implemented in other classrooms and discussed the design and content of the curriculum. Teachers who implemented collaborative group work learned about the goal of the intervention, the research and theory supporting collaborative reasoning, how to facilitate CR discussions, and effective strategies for promoting group work. They saw videos of the wolf unit being taught using collaborative groups. Teachers who implemented whole-class direct instruction learned about the research and theory supporting explicit teaching of concepts and strategies and effective teaching strategies for direct instruction. They saw videos of the wolf unit being taught using direction instruction.

The wolf unit occupied about an hour a day, four or five days a week, for six weeks. Throughout the unit, in both CG and DI classrooms, a research assistant was onsite working with teachers to ensure fidelity of implementation, video-record lessons, make field notes, and administer tests. Wait-listed control classes completed pretests and posttests but continued to receive regular language arts instruction during the intervention period. In the following semester, control teachers were invited to a workshop and offered the opportunity to implement the wolf unit in their classes.

Outcome Measures

After the Wolf Reintroduction and Management Unit, extensive information was gathered to evaluate outcomes in CG, DI, and control classrooms. Students completed two tests assessing *unit mastery*: a 50-minute individually written essay in which they explained their personal decision about whether the pack of wolves should be eradicated and a 100-item sentence verification test that addressed the main concepts and information in the wolf unit. *Near transfer* was evaluated in a one-to-one interview about an analogue to the wolf question, whether whaling should be allowed. *Far transfer* tasks consisted of an oral narrative prompted by the wordless

picture book, *Frog, Where Are You*, and an independently written essay involving a decision about a moral and practical dilemma posed by *The Pine Wood Derby* story. *Near* and *far* are customary although informal ways of referring to the degree of similarity between the training task and the transfer task. The distinction is informal because, as Salomon and Perkins (1989) have noted, there is no satisfactory way to scale all of the features of tasks and contexts that contribute to similarity.

The current article reports analyses of the far transfer task of reflective essay writing based on *The Pine Wood Derby*. In this story, Thomas is portrayed as a poor and unpopular kid who has never won anything. He finally gets his chance and wins the championship in the Pinewood Derby model car competition. But his older brother helps him make the car, which is against the rule of working independently. He tells this secret to his friend, Jack, and asks him to keep his secret. Jack is disappointed at Thomas's dishonesty; on the other hand, he is glad his poor friend has finally won something. The decision that students were asked to make is whether or not Jack should tell on Thomas. Students had 50 minutes to compose the essay; most of them finished before the time limit.

Essays prompted by *The Pine Wood Derby* have provided the outcome measure in several previous CR studies (e.g., Kim et al., 2011; Reznitskaya et al., 2001); however, in these studies, writing the essay was a near transfer task because the preceding collaborative discussions were based on similar moral and practical dilemmas. In the current project, writing the essay becomes a far transfer task because *The Pine Wood Derby* has little in common with the wolf unit. Thus, comparisons of performance on this task by children in CG, DI, and control classrooms provides a strong basis for determining which children have acquired generalizable competence in decision making.

Coding Children's Reflective Essays

Three aspects of the Jack and Thomas essays were coded: (a) whether two or more sides of a dilemma were articulated; that is, the essay expressed costs as well as benefits of decisions. For instance, if a child thought that Jack should tell on Thomas because of his dishonesty, did he or she also mention possible negative effects of this decision, such as Jack would be called a tattletale by his classmates or might lose Thomas as a friend? (b) Ability to generate different types of reasons. These reasons were further broken down into nine moral principles and 10 practical considerations, as enumerated in Table 1. These reasons have previously been shown to exhaust the reasons children consider in the *Pinewood Derby* essay (Zhang et al., 2013). And (c) explicitness in weighing different options; for example, stating that being called a tattletale is trivial compared to protecting fairness.

Children's essays were exported into the NVivo8 qualitative research software (QSR, 2008) for coding. Raters were blind to whether students

Table 1

Definitions and Examples of Moral Principles and Practical Considerations

Moral Principles		Practical Considerations	
Category	Definition or Example	Category	Definition or Example
Honesty	Thomas did not make the car by himself. He broke a rule because he got assistance from his brother.	Jack will become an accomplice.	Keeping the secret for Thomas makes Jack lie to the whole class; therefore, he would become an accomplice.
Fairness	Thomas's use of assistance was not fair to other students in the class who made cars by themselves.	Not telling on Thomas will do harm to him.	Jack's tolerance for Thomas's misdeed will make him behave even more badly in the future.
Empathy	Thomas was a poor kid, and he had never won anything. Jack's mother told him to be nice to people who are less fortunate.	Jack will get into trouble.	Thomas might beat up Jack if he tells the secret.
Friendship	Jack and Thomas are friends. Friends help each other keep secrets, or friends help each other overcome weakness.	The teacher will not believe Jack.	Mr. Howard, the teacher of the class, might not believe Jack if he told him that Thomas cheated.
Promise	Thomas asked Jack to keep this secret for him.	History will not rewrite.	What happened has happened. There is nothing we can do to change the fact that Thomas got the trophy.
Trust	Thomas trusted Jack, that's why he told the secret to Jack instead of somebody else.	Jack will be rewarded for telling the truth.	Jack may get an award if he tells the secret.
Golden Rule	Treat others as one would like others to treat oneself; do not treat others in ways that one would not like to be treated.	Thomas will get mad.	Thomas will be angry if he learns that Jack told his secret to the teacher.
Common good	Jack should make a decision that is beneficial for all or most students.	Jack will be seen as a sore loser.	People would think that Jack tells because he feels jealous.
Don't tattle	It is wrong to betray friends to authorities.	Thomas is mean.	Thomas pushes kids for no reason and calls them names.
		Winning makes Thomas a better boy.	By winning the prize, Thomas feels more confident of himself; he will do better in the future.

were from CG, DI, or control classrooms. They had no information about students' gender, ethnicity, reading comprehension scores, or other personal characteristics. Both dilemma and weighing were coded as binary variables, with 1 indicating present and 0 representing absent. A second rater independently coded 20% of the essays. Interrater reliability was satisfactory (Cohen's kappa = .87 for dilemma, .81 for weighing).

For the content analysis of types of reasons, each of the nine moral principles and 10 practical considerations was coded as present (1) or absent (0). Number of reasons is the sum of the number of moral principles and practical considerations. This measure represents the range of different reasons children appealed to when making decisions. A second rater independently coded 20% of the essays. Interrater reliability was satisfactory (Cohen's kappa = .75 for moral principles, .76 for practical considerations).

Results

We conducted three sets of analyses of children's decision-making competence, one for each of the three aspects of sound decision making: recognizing two sides of a dilemma, comprehensiveness of reasons, and weighing reasons according to importance. Excerpts from children's essays, with spelling corrected (to enable keyword searches) but otherwise unedited, are presented in Table 2 to illustrate each aspect of decision making. No one will fail to notice that the excerpts contain grammatical mistakes, awkward wording, and difficult to grasp ideas. It needs to be reiterated that as a group, the children who participated in this study were well below the national average in academic attainment, represented in this study by a standardized measure of reading comprehension. Moreover, the essays were the children's first drafts, and to do their best work, children need the opportunity to edit and improve their writing in several drafts (Graves, 1994). Table 3 presents the means and standard deviations of the three measures of decision making by instructional condition.

Considering Two Sides of a Dilemma

Children who managed to recognize more than one side of a dilemma in making a decision about whether Jack should tell on Thomas were either able to display a rationale for each side or to see undesirable implications of their own preferred decision (for illustrations, see Table 2).

Since individuals were nested within classrooms, a two-level logistic regression model was constructed to predict consideration of more than one side of a dilemma. Individual-level predictors were gender, ethnicity, Gates-MacGinitie reading comprehension (classroom-mean centered), basic English proficiency as represented by time to name common objects (also classroom-mean centered), and condition (CG, DI, control). Classroom-level predictors were classroom mean reading comprehension and classroom

Table 2

Coding Children's Essays: Recognition of Two sides of a Dilemma, Comprehensiveness of Reasons, and Weighing Importance of Reasons

Decision Feature	Coding Criteria	Example	Explanation
Considering two sides of a dilemma	Displaying a rationale for each side of the decision or undesirable implications of the preferred decision	I think that Jack should tell on Thomas because all the other kids worked harder than the other kids. Thomas's brother made sure that everything was perfect so Thomas can win. Thomas didn't do it by himself. He just put the paint and the stickers. . . . If Jack tells Mr. Howard that Thomas cheated. Nobody will believe him cause people thinks that his a tattletale. No one will want to be friends with Jack because he told on Thomas and Jack promised to Thomas that he wouldn't tell on the fixing.	In this excerpt, the child recognized the unfavorable consequences (e.g., being a tattletale, breaking the promise) associated with his or her preferred decision.
Comprehensiveness of reasons	Considering different types of reasons, including moral principles and practical considerations (see details in Table 1).	If I was Jack I wouldn't had told on Thomas. I wouldn't because Thomas probably was made fun of half of his life so that's why he did the things he did. Even though Thomas had won unfairly it was good that more kids admired him and started talking. If more people would pay attention to him he would not be so mean and stressed. His whole life was maybe stressful. So he felt he had to stress other kids. I also think Thomas should thank Jack for not telling Mr. Howard. But also, Thomas should thank Jack and try to become friends with him so he won't be so mean to anyone. I think Thomas would be nice if he had friends at school. To me he seemed kind of nice and a harmless child. But he just had a lot of stress in him. So he had to take his anger out on someone else other than his family. But overall, Jack shouldn't have told anyone about Thomas cheating. He probably never ever won a race or game at school or anything. So he was finally admired because he won a Pinewood Derby contest and a trophy. Hopefully he gets more friends. But Jack was a nice friend for not telling and keeping a promise to Thomas. I hope Jack and Thomas become friends after all.	This excerpt contained five different types of reasons. Four different moral principles were mentioned by the child: honesty (e.g., Thomas cheating), fairness (e.g., Thomas won unfairly), empathy (e.g., He probably never ever won a game or race at school), and friendship (e.g., Jack is a nice friend for not telling on Thomas). The practical consideration of making Thomas a better boy (e.g., Thomas won't be so mean to anyone/he would be nice to everyone if he has friends) was also included in the essay.

(continued)

Table 2 **(continued)**

Decision Feature	Coding Criteria	Example	Explanation
Weighing the importance of reasons	Explicitly comparing two reasons, using phrases such as *more important than, less important than*, and so on, or phrases that marked a reason as important or unimportant without comparing it with another reason, such as *it is for a good cause, just do the right thing, who cares being the tattletale*	Jack shouldn't tell on Thomas because he was the only one who was nice to him. Everybody called him a name. Also the only thing the he didn't do was glue the wheels on but he did everything else. But at the same time he should tell on him because Jack had work really hard and he did not need help from nobody. . . . Jack said that would tell nobody that he didn't work on it by himself. But I think if Jack tells he is going to lose his friend that was nice to him. Also he didn't make fun of him. *But he was he friend the whole time that nobody was his friend and that is more important than lose a friend then lose a game and nobody got mad about the game*. So he shouldn't tell on him because he was to him and didn't tries to be mean to you and that he was when you didn't have no friend and *friend more important they a game*. Also never call him small like all of the other kids was and no one was tries to be mean to him but the other kids was being mean the him. So he shouldn't tell on him.	In this excerpt, the child endorsed the decision not to tell on Thomas, in spite of his dishonesty, by giving greater weight to friendship than winning the contest.

mean basic English proficiency. A series of models were fitted. The initial model contained all main effects and their interactions. Nonsignificant terms were dropped one at a time if there was not a significant difference between the model with the term and the model without it, as indicated by the change in χ^2 of the two models. None of the interaction terms (e.g., condition × gender, condition × reading comprehension) were significant, and retaining them led to poorer model fits; thus, interactions were excluded from subsequent models. Since neither individual nor classroom mean basic English proficiency significantly predicted the outcome measure, these terms were removed. Although not significant, ethnicity was retained; results do not change when ethnicity is removed from the final model shown in the following.

Level 1: Individual-level model is

$$\text{Dilemma}_{ij} = \beta_{0j} + \beta_{1j}(\text{gender})_j + \beta_{2j}(\text{ethnicity})_j + \beta_{3j}(\text{individual reading comprehension})_j + \beta_{4j}(\text{condition}).$$

Level 2: Classroom-level model is

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{classroom mean reading comprehension})_j + U_{0j}.$$

As shown in Table 4, condition and gender were significant factors predicting consideration of more than one side of a dilemma. After controlling for other factors, including any effects associated with classroom, CG students were significantly better than their counterparts in DI classrooms, $\beta = .45$, $SE = .23$, odds ratio = 1.57, 95% CI [1.02, 2.43], $t(683) = 2.03$, $p < .05$, in recognizing different sides of a dilemma; however, no significant differences were found between CG and control classrooms, $\beta = .25$, $SE = .23$, odds ratio = 1.23, 95% CI [.82, 2.00], $t(683) = 1.07$, $p = .29$, or between DI and control classrooms, $\beta = -.19$, $SE = .23$, odds ratio = .83, 95% CI [.53, 1.30], $t(683) = -.91$, $p = .37$. Girls were better than boys in perceiving the dilemma, $\beta = .89$, $SE = .17$, odds ratio = 2.47, 95% CI [1.77, 6.45], $t(683) = 5.33$, $p < .001$.

Comprehensiveness of Reasons

A deliberate and thoughtful decision-making process is marked not only by an individual's ability to attend to both sides of an issue but also by the variety of reasons that he or she can consider (see Table 1 for the classification of reasons).

Table 5 presents the two-level Poisson regression analysis of the total number of distinct reasons, including both moral principles and practical considerations. Following the same model fitting and selection procedures described previously, the final models included all main effects other than individual or classroom mean basic English proficiency. The results indicate

Table 3
Means (*SD*s) by Condition for Recognizing Two Sides of the Dilemma, Comprehensiveness of Reasons, and Weighing Importance of Reasons

	Condition		
Decision Feature	CG (N = 245)	DI (N = 249)	Control (N = 233)
Two sides of dilemma	0.40 (.49)	0.30 (.46)	0.33 (.47)
Comprehensiveness of reasons	2.75 (1.39)	2.27 (1.21)	2.27 (1.11)
Weighing importance	0.22 (.41)	0.11 (.31)	0.13 (.34)

Note. CG = collaborative groups; DI = direct instruction.

that CG students appealed to a significantly greater number of distinct reasons in their essays than DI students, $\beta = .17$, $SE = .07$, $t(683) = 2.47$, $p < .05$, or control students, $\beta = .16$, $SE = .07$, $t(683) = 2.23$, $p < .05$, whereas students from DI and control classrooms were indistinguishable from each other, $\beta = -.01$, $SE = .07$, $t(683) = -.15$, $p = .88$. Glass's Δ (Glass, 1976), which is suited to effect size comparisons involving skewed outcomes measures, was $\Delta = .40$ for CG versus DI and $\Delta = .43$ for CG versus control. Girls included significantly more reasons in their essays than boys, $\beta = .16$, $SE = .05$, $t(683) = 3.32$, $p < .001$. The results were the same when ethnicity was removed from the final model.

Weighing the Importance of Reasons

For students who were able to present two sides of a dilemma in their essays, we further investigated whether their essays contained the decision-making criterion of weighing. Relying on the same model fitting and selection procedures as before, a two-level logistic regression of the likelihood of weighing competing arguments was fitted. Table 4 shows that students in CG classrooms performed significantly better than students in either DI classrooms, $\beta = .80$, $SE = .27$, odds ratio = 2.22, 95% CI [1.32, 3.73], $t(683) = 3.03$, $p < .01$, or control classrooms, $\beta = .54$, $SE = .27$, odds ratio = 1.72, 95%, CI [1.02, 2.89], $t(683) = 2.19$, $p < .05$. DI students did not differ from their counterparts in control classrooms, $\beta = -.26$, $SE = .27$, odds ratio = .77, 95% CI [.44, 1.37], $t(683) = -.89$, $p = .37$. Girls were more likely to weigh reasons than boys, $\beta = .97$, $SE = .24$, odds ratio = 2.63, 95% CI [1.66, 4.17], $t(683) = 4.10$, $p < .001$. Children with higher reading comprehension had a greater tendency to weigh reasons than less proficient children, $\beta = .03$, $SE = .01$, $t(683) = 2.10$, $p < .05$. Ethnicity did not make a significant difference.

General Discussion

The major finding of the present study is that students in collaborative group work classrooms wrote essays that were significantly better than the

Table 4

Two-Level Logistic Regression Models Predicting Recognition of Two Sides of the Dilemma and Weighing Importance of Reasons as a Function of Condition, Gender, Ethnicity, and Reading Comprehension

	Two Sides of Dilemma				Weighing Importance of Reasons			
			95% CI				95% CI	
Predictors	Coefficient (*SE*)	Odds ratio	LL	UL	Coefficient (*SE*)	Odds ratio	LL	UL
Intercept	–1.59*** (0.38)				–2.95*** (0.47)			
Level 1: Individual level								
Gender	0.89*** (0.17)	2.47	1.77	6.45	0.97*** (.24)	2.63	1.66	4.17
Ethnicity	0.03 (0.18)	0.96	0.68	1.36	0.03 (0.22)	0.97	0.63	1.49
Reading comprehension	0.01 (0.01)				0.03* (0.01)			
Condition								
CG vs. DI	0.45* (0.23)	1.57	1.02	2.43	0.80** (0.27)	2.22	1.32	3.73
DI vs. control	–0.19 (0.23)	0.83	0.53	1.30	–0.26 (0.27)	0.77	0.44	1.37
CG vs. control	0.25 (0.23)	1.23	0.82	2.00	0.54* (0.27)	1.72	1.02	2.89
Level 2: Classroom level								
Class mean reading comprehension	0.03 (0.02)				0.03 (0.03)			

Note. Gender coded girl = 1, boy = 0. Ethnicity coded Hispanic = 1, others = 0. CG = collaborative groups; DI = direct instruction.
*$p < .05$. **$p < .01$. ***$p < .001$.

Table 5

Two-Level Poisson Regression Model of Comprehensiveness of Reasons as a Function of Condition, Gender, Ethnicity, and Reading Comprehension

	Comprehensiveness of Reasons	
Predictors	Coefficient (*SE*)	Glass's Δ
Intercept	.57*** (.11)	
Level 1: Individual level		
Gender	.16*** (.05)	.32
Ethnicity	.09 (.06)	.19
Reading comprehension	.00 (.00)	
Condition		
CG vs. DI	.17* (.07)	.40
DI vs. control	–.01 (.07)	.00
CG vs. control	.16* (.07)	.43
Level 2: Classroom level		
Class mean reading comprehension	.01 (.01)	

Note. Gender coded girl = 1, boy = 0. Ethnicity coded Hispanic = 1, others = 0. CG = collaborative groups; DI = direct instruction.
*$p < .05$. ***$p < .001$.

essays of students from direct instruction classrooms in each of three aspects of decision making—recognizing more than one side of a dilemma, considering a range of reasons, and weighing the importance of reasons. In contrast, DI students performed no better than uninstructed control students.

Comparing the CG and control conditions, CG students trended better in terms of recognizing two sides of a dilemma; however, the difference was not statistically significant, which is inconsistent with findings from previous studies (e.g., Dong et al., 2008, 2009; Reznitskaya et al., 2001). A possible reason for the discrepancy is that the *Pinewood Derby* story was a near transfer task in prior studies but a far transfer task in the current research. In the previous studies, students in the collaborative reasoning condition discussed stories that contained similar moral or practical dilemmas, which might have increased their ability to recognize both sides of the dilemma in the *Pinewood Derby* story. CG students were significantly better than control students in the other two aspects of decision making, considering a variety of reasons and weighing the importance of reasons. Taken together, these findings demonstrate far transfer for students who experienced collaborative interaction, inasmuch as wolf reintroduction and management, which CG and DI students had studied for six weeks, has few if any surface features in common with the transfer task—whether or not a boy should tell the teacher that his friend has cheated in a contest building and racing model cars.

Performance on the Wolf Reintroduction and Management Unit mastery tests provides a benchmark for interpreting the transfer effects since that which is not learned cannot be transferred. After the intervention, students completed two unit mastery tests: a 100-item sentence verification test that addressed the main concepts and information in the wolf unit and a 50-minute individually written essay in which they explained their personal decision about whether the pack of wolves should be eradicated. Analysis of performance on the sentence verification test indicated that DI students grasped significantly more concepts and information from the wolf unit than CG students or control students (for details, see Morris, Anderson, & Collaborative Reasoning Research Group, 2013). Morris and colleagues (2013) also examined argument elements in the essays children wrote to explain and justify their decisions about whether the pack of wolves should be eradicated. Using a coding scheme somewhat different from the one employed here, based on an argument-counterargument-rebuttal structure, Morris et al. (2013) found that CG and DI students were equally good at considering counterarguments and rebutting counterarguments as well as producing elaborated reasons based on the concepts and information from the wolf unit. Thus, DI students were able to use knowledge acquired from the wolf unit in an essay that was a recapitulation of the unit.

However, in an essay about a novel situation, the dilemma faced by the boy in the *Pinewood Derby* story, DI students were unable to make use of knowledge that they had apparently acquired. The contrast of DI students' performance on the unit mastery tests and the far transfer task seems to indicate that DI students had acquired "inert ideas," ideas that had been "received into the mind" but which could not be "utilized . . . or thrown into fresh combinations" (Whitehead, 1929/1957).

Direct instruction students and students who participated in collaborative groups studied the same curriculum, so the differences between DI and CG on the far transfer task in the three aspects of decision making must arise from the contrasting conditions of learning. We theorize that the essential difference between direct instruction and collaborative group work lies in what students learn about the "self as agent and others as audience" (Morris et al., 2013). In DI classrooms, teachers do most of the initiating and evaluating, leaving students with circumscribed opportunities for extended reasoning and, we suppose, little cause to believe that their own thinking counts for much in making a decision about the wolves. In a collaborative learning environment, students have the chance to become active arguers and decision makers who engage in chains of reasoning themselves (Ma et al., 2015). CG students are more likely to understand that, like themselves, other people may need help in understanding the implications of ideas, so they work together with others to clarify issues and look at the issues from different perspectives. CG students, in other words, have

a stronger sense that they ought to test their ideas against the competing ideas of others, either to persuade or to be persuaded.

Explaining the improved performance of CG students in the essay about the model car competition necessarily involves the assumption that CG students were able to see deep relational correspondences (Chi & VanLehn, 2012; Hummel & Holyoak, 2003) between the wolf management problem and the question about cheating in the model car race because the two situations do not share common surface features. The set of relational correspondences in common between these and other situations that call for decision making may be said to comprise a decision-making *schema* (Anderson & Pearson, 1984; Reznitskaya et al., 2009; Rumelhart, 1980), or overall sense of when and how to engage in a deliberate decision-making process. We put the focus on process rather than outcome because it is incorrect to equate a sound decision-making process with a favorable outcome, as unexpected or uncontrollable events may affect an outcome. However, in the long term, knowing how to make reasonable decisions should lead to an increased chance of obtaining good outcomes compared to a thoughtless or imprudent decision-making process (Clemen & Hampton, 1994).

The Wolf Reintroduction and Management Unit afforded an opportunity to learn something about making decisions about complex and controversial issues. CG students were positioned to take advantage of this opportunity. DI students were not. DI students were positioned as bystanders looking on as others, described in the instructional materials and by the teacher, considered the decision about the wolves. Rumelhart (1980) likened a schema to a script for a play, explaining that,

> Just as a play has characters that can be played by different actors at different times without changing the essential nature of the play, so a schema has variables that can be associated with (bound to) different aspects of the environment on different instantiations of the schema. (p. 35)

The CG student could readily bind to the role of decision maker, the DI student not so readily or not at all.

Daily engagement in dialogue with others was expected to raise CG students' awareness of varied perspectives on the decision about wolves. Increasing awareness might be realized through the *snowball* process: When one child introduces a new perspective or another way of reasoning, it often spreads to other children. Anderson and colleagues (2001) found that once a child utilized a new argument strategy in a small-group collaborative discussion—for instance, linking propositions with *but* as a way to show disagreement and introduce a counterargument—other discussants immediately picked up the strategy and began using it. The consequence was increasing use of worthwhile argument strategies over time and among

increasing numbers of discussion participants. Lin et al. (2012) reported that strategies as complicated as arguing by analogy snowball during collaborative discussions.

CG students appealed to a significantly greater number of different moral principles and practical considerations in their essays. Plausibly, many of these were generated analogically from arguments that occurred during the wolf unit discussions. For example, the principle of fairness would have been invoked in an argument for wolves' right to live or an argument about the infringement of the interests of tourism-dependent business owners if wolves were eradicated. The Golden Rule entails taking another's perspective, like putting oneself into the position of wolves. Based on the snowball process, if one child brings up a new perspective, other children in the group may pick it up, elaborate, and extend it. Eventually, children internalize these moral principles and develop greater insight into the circumstances in which they apply, allowing them to flexibly use the principles when they encounter a novel situation. The following is an example in which the Golden Rule was applied in a collaborative discussion by Albert. The form is echoed by Sandra, but she takes the other side, saying that she and her family would not like it if a wolf killed their pet.

> [00:25:22.00] Albert: Yeah, but wolves are like humans like they have like like if you were a wolf how would you feel if somebody came and killed you? . . . You don't like it. [00:25:32.01]
>
> [00:25:34.20] Sandra: But we like we have an animal and if it will kill it we would not (neither) like it. [00:25:41.14]

When making a decision, the weighing of options builds on the recognition of the competing sides of a dilemma and awareness of various reasons that can be advanced on each side. Piaget (1947/1976; see also, Johnson & Johnson, 2009) maintained that real learning does not occur until children reach a state of disequilibrium when newly encountered information contradicts their previous beliefs. During collaborative discussions, when a conflict between a student's belief and others' opinions arose, there was a need for them to reconcile the difference, and this is likely to be the time when weighing occurred. Students in CG classrooms had many opportunities to compare competing claims and hear others compare competing claims, which would introduce them to the process of weighing.

Initial analyses of classroom talk during the wolf unit are consistent with the explanation that differences in the dialogue in CG and DI classrooms figured in the differential development of decision making of CG and DI students (Ma et al., 2015; Morris et al., 2013). Morris and colleagues (2013) analyzed the frequency of use of coordinating conjunctions in video clips systematically sampled from the approximately 500 hours of lesson videos recorded in CG and DI classrooms. Conjunctions showing how ideas are

connected are a low-inference indicator of reasoning (Lin et al., 2015). *Because* and *so* can mark causal connections, *if* and *then* can indicate temporal and logical relations, while the contrastive *but* can be used to introduce a counterargument, as is illustrated in Table 2. CG students' rate of use of the coordinating conjunctions *because, so, if, then, and*, and *but* was more than four times higher than the rate of DI students. In contrast, interestingly, DI teachers used the set of conjunctions at over twice the rate of CG teachers. The picture that emerges of the DI classroom is one in which the teacher does most of the reasoning, with students filling in small pieces when requested, whereas in CG classrooms, students do most of the reasoning, with the teacher occasionally redirecting when students go off track.

As a specific indicator of whether students considered themselves to be active decision-making agents, Morris and colleagues (2013) searched the video clips from the CG and DI classrooms for the performative verb phrases *I think* and *I know*. CG students used these phrases at twice the rate of DI students.

Girls performed significantly better than boys in the three aspects of decision making: recognizing a dilemma, appealing to a range of reasons, and weighing reasons on the two sides, perhaps related to the general finding that girls are better writers than boys (e.g., National Center for Education Statistics, 2011). Gender did not interact with condition, meaning that CG and DI had a comparable influence on boys and girls.

The present study shows the value of an embedded issue-based socioscientific decision-making curriculum in helping young students become aware of how to make reasonable and responsible decisions. Importantly, the curriculum was successful only when students participated in collaborative groups. Students who received teacher-led direct instruction did no better than uninstructed controls. The study is an advance over previous studies because it demonstrates that students who studied the decision-marking curriculum in collaborative groups were able to transfer what they had learned to a very different task. This implies that they had acquired generalizable skills and dispositions important for decision making.

An educational policy implication of our findings is that if children are to start on the path toward becoming thoughtful decision makers, they need more time during the school day for active reasoning about significant issues. More time for active reasoning about significant issues is critical for children of color because it is one key to "affirmative development of intellective competence and academic ability" (Gordon, 2014, p. 26). But since the passage of No Child Left Behind, the nation's schools have been moving in the opposite direction. The race to the bottom has been swift in schools with large enrollments of minority and low-income children, who now spend most of the school day on arithmetic exercises and simple reading strategies. Science, social studies, art, and music have been stripped from

the curriculum, and recess has been eliminated to make more room for basic skills instruction. Collaborative learning in small groups has given way to teacher-directed whole-class instruction and individual seatwork, which are assumed to be more efficient for transmission of basic skills. The Common Core State Standards were supposed to represent a change of direction. The standards do stress reasoning and other higher-order skills, but as put into practice, the Common Core is perpetuating a test-driven accountability system and a teacher-dominated approach to discussion that are inconsistent with the development of reasoning.

Note

The research reported in this article was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A080347 to the University of Illinois at Urbana-Champaign, Richard C. Anderson, Principal Investigator. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. The authors are pleased to acknowledge the contributions of Erjing Cui, Nikisha Blackmon, Xiaoying Wu, and Aini Marina Ma'rof.

References

Anderson, R. C., Chinn, C., Waggoner, M., & Nguyen, K. (1998). Intellectually stimulating story discussions. In J. Osborn & F. Lehr (Eds.), *Literacy for all: Issues in teaching and learning* (pp. 170–186). New York, NY: Guilford Press.

Anderson, R. C., Nguyen-Jahiel, K., McNurlen, B., Archodidou, A., Kim, S.-Y., Reznitskaya, A., . . . Gilbert, L. (2001). The snowball phenomenon: Spread of ways of talking and ways of thinking across groups of children. *Cognition and Instruction*, *19*, 1–46. doi:10.1207/S1532690XCI1901_1

Anderson, R. C., & Pearson, P. D. (1984). A schema-thematic view of basic processes in reading comprehension. In P. D. Pearson, R. Barr, M. L. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 255–291). New York, NY: Longman.

Asterhan, C. S. C., & Schwarz, B. B. (2007). The effects of monological and dialogical argumentation on concept learning in evolutionary theory. *Journal of Educational Psychology*, *99*(3), 626–639. doi:10.1037/0022-0663.99.3.626

Asterhan, C. S. C., & Schwarz, B. B. (2009). Argumentation and explanation in conceptual change: Indications from protocol analyses of peer-to-peer dialog. *Cognitive Science*, *33*(3), 374–400. doi:10.1111/j.1551-6709.2009.01017.x

Bakhtin, M. M. (1981). *The dialogic imagination: Four essays by M.M. Bakhtin*. Austin, TX: University of Texas Press.

Beyth-Marom, R., Fischhoff, B., Quadrel, M. J., & Furby, L. (1991). Teaching decision making to adolescents: A critical review. In J. Baron & R. V. Brown (Eds.), *Teaching decision making to adolescents* (pp. 19–59). Hillsdale, NJ: Lawrence Erlbaum Associates.

Braund, M., Lubben, F., Scholtz, Z., Sadeck, M., & Hodges, M. (2007). Comparing the effect of scientific and socio-scientific argumentation tasks: Lessons from South Africa. *School Science Review*, *88*(324), 67–76.

Byrnes, J. P. (2002). The development of decision-making. *Journal of Adolescent Health*, *31*(6), 208–215.

Chapman, G. B., & Sonnenberg, F. A. (Eds.). (2003). *Decision making in health care: Theory, psychology, and applications*. Cambridge, UK: Cambridge University Press.

Chi, M. T. H., & VanLehn, K. A. (2012). Seeing deep structure from the interactions of surface features. *Educational Psychologist*, *47*(3), 177–188.

Clemen, R. T., & Hampton, H. (1994). *Cooperative learning and decision making*. Eugene, OR: Decision Research.

Dillenbourg, P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and computational approaches* (pp. 1–19). Oxford, UK: Elsevier.

Dong, T., Anderson, R. C., Kim, I., & Li, Y. (2008). Collaborative reasoning in China and Korea. *Reading Research Quarterly*, *43*(4), 400–424.

Dong, T., Anderson, R. C., Lin, T., & Wu, X. (2009). Concurrent student-managed discussions in a large class. *International Journal of Educational Research*, *48*(5), 352–367.

Dye, T. R. (1992). *Understanding public policy*. Englewood Cliffs, NJ: Prentice Hall.

Engle, S. H. (1960). Decision making: The heart of social studies instruction. *Social Education*, *24*(7), 301–304.

Fawcett, L. M., & Garton, A. F. (2005). The effect of peer collaboration on children's problem-solving ability. *The British Journal of Educational Psychology*, *75*, 157–169. doi:10.1348/000709904X23411

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, *62*, 451–82. doi:10.1146/annurev-psych-120709-145346

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, *5*, 3–8.

Gordon, E. W. (2014). *Pedagogical imagination: Volume II: Using the master's tools to inform conceptual leadership, engaged scholarship and social action*. Chicago, IL: Third World Press.

Graves, D. H. (1994). *A fresh look at writing*. Portsmouth, NH: Heinemann

Halpern-Felsher, B. L., & Cauffman, E. (2001). Costs and benefits of a decision: Decision-making competence in adolescents and adults. *Journal of Applied Developmental Psychology*, *22*(3), 257–273.

Hibbard, J. H., Slovic, P., & Jewett, J. J. (1997). Informing consumer decisions in health care: Implications from decision-making research. *Milbank Quarterly*, *75*(3), 395–414.

Howlett, M. (2009). Policy advice in multi-level governance systems: Sub-national policy analysts and analysis. *International Review of Public Administration*, *13*(3), 1–16.

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review*, *104*, 427–466.

Johnson, D. W., & Johnson, R. T. (2009). Energizing learning: The instructional power of conflict. *Educational Researcher*, *38*, 37–51.

Jonassen, D. H. (2012). Designing for decision making. *Educational Technology Research and Development*, *60*(2), 341–359. doi:10.1007/s11423-011-9230-5

Kim, I., Anderson, R. C., Miller, B., Jeong, J., & Swim, T. (2011). Influence of cultural norms and collaborative discussions on children's reflective essays. *Discourse Processes*, *48*(7), 501–528. doi:10.1080/0163853X.2011.606098

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*, 75–86.

Klaczynski, P. A. (2001). Analytic and heuristic processing influences on adolescent reasoning and decision-making. *Child Development*, *72*, 844–861. doi:10.1111/1467-8624.00319

Klaczynski, P. A., Gordon, D. H., & Fauth, J. (1997). Goal-oriented critical reasoning and individual differences in critical reasoning biases. *Journal of Educational Psychology*, *89*, 470–485. doi:10.1037/0022-0663.89.3.470

Kuhn, D. (1992). Thinking as argument. *Harvard Educational Review*, *62*(2), 155–179.

Kuhn, D., & Udell, W. (2003). The development of argument skills. *Child Development*, *74*(5), 1245–1260.

Lee, Y. C. (2007). Developing decision-making skills for socio-scientific issues. *Journal of Biological Education*, *41*(4), 170–178.

Lee, Y. C., & Grace, M. (2010). Students' reasoning processes in making decisions about an authentic, local socio-scientific issue: Bat conservation. *Journal of Biological Education*, *44*(4), 156–165. doi:10.1080/00219266.2010.9656216

Lin, T.-J., Anderson, R. C., Hummel, J. E., Jadallah, M., Miller, B. W., Nguyen-Jahiel, K., . . . Dong, T. (2012). Children's use of analogy during collaborative reasoning. *Child Development*, *83*, 1429–1443. doi:10.1111/j.1467-8624.2012.01784.x

Lin, T.-J., Jadallah, M., Anderson, R. C., Nguyen-Jahiel, K., Kim, I.-H., Kuo, L.-J., . . . Li, Y. (2015). Social influences on the development of relational thinking during small-group discussions. *Contemporary Educational Psychology*, *41*, 83–97. doi:10.1016/j.cedpsych.2014.12.004

Ma, S., Anderson, R. C., Lin, T.-J., Zhang, J., Morris, J., Nguyen-Jahiel, K. T., . . . Latawiec, B. (2015). *Influence of collaborative group work and direct instruction on English language learners' oral narratives*. Champaign, IL: Center for the Study of Reading.

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L.G. (2000). *Gates-MacGinitie reading tests (4th ed.), Level 4, Form S*. Itasca, IL: Riverside Publishing.

Mann, L., Harmoni, R. O. S., Power, C., Beswick, G., & Ormond, C. (1988). Effectiveness of the GOFER course in decision making for high school students. *Journal of Behavioral Decision Making*, *1*, 159–168. doi:10.1002/bdm.3960010304

Mercer, N. (1996). The quality of talk in children's collaborative activity in the classroom. *Learning and Instruction*, *6*, 359–378.

Mercer, N. (1999). Children's talk and the development of reasoning in the classroom. *British Educational Research Journal*, *25*(1), 95–111.

Mercer, N., Dawes, R., Wegerif, R., & Sams, C. (2004). Reasoning as a scientist: Ways of helping children to use language to learn science. *British Educational Research Journal*, *30*, 367–385.

Morris, J., & Anderson, R. C., & the Collaborative Reasoning Research Group. (2013). *Instructional discourse and argumentative writing*. Champaign, IL: Center for the Study of Reading.

National Center for Education Statistics. (2012). *The nation's report card: Writing 2011* (NCES 2012–470). Washington, DC: Institute of Education Sciences, U.S. Department of Education.

Newman, F. M., & Oliver, D. W. (1970). *Clarifying public controversy*. Boston, MA: Little Brown and Company.

Newman, F. M. (1990). Higher order thinking in teaching social studies: A rationale for the assessment of classroom thoughtfulness. *Journal of Curriculum Studies*, *22*(1), 41–56.

Nystrand, M., Wu, L., Gamoran, A., Zeiser, S., & Long, D. (2003). Questions in time: Investigating the structure and dynamics of unfolding classroom discourse. *Discourse Processes*, *35*, 135–198.

Piaget, J. (1932). *The moral judgment of the child*. London: Routledge & Kegan Paul.

Piaget, J. (1976). *The psychology of intelligence*. New York, NY: Littlefield. (Original work published 1947)

QSR. (2008). QSR NVivo8 [Computer software]. Vicky, Australia: Qualitative Solutions and Research.

Raiffa, H. (1968). *Decision analysis: Introductory lectures on choices under uncertainty*. Reading, MA: Addison-Wesley.

Ratcliffe, M. (1997). Pupil decision-making about socio-scientific issues within the science curriculum. *International Journal of Science Education*, *19*(2), 167–182. doi:10.1080/0950069970190203

Regenwetter, M., Grofman, B., Popova, A., Messner, W., Davis-Stober, C. P., & Cavagnaro, D. R. (2009). Behavioural social choice: A status report. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1518), 833–843. doi:10.1098/rstb.2008.0259

Reznitskaya, A., Anderson, R. C., & Kuo, L. (2007). Teaching and learning argumentation. *The Elementary School Journal*, *107*(5), 449–472.

Reznitskaya, A., Anderson, R. C., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A., & Kim, S. (2001). Influence of oral discussion on written argument. *Discourse Processes*, *32*(2–3), 155–175. doi:10.1080/0163853X.2001.9651596

Reznitskaya, A., Kuo, L., Clark, A., Miller, B., Jadallah, M., Anderson, R. C., & Nguyen-Jahiel, K. (2009). Collaborative reasoning: A dialogic approach to group discussions. *Cambridge Journal of Education*, *39*(1), 29–48. doi:10.1080/03057640802701952

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York, NY: Oxford University Press.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 38–58). Hillsdale, NJ: Lawrence Erlbaum Associates.

Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanisms of a neglected phenomenon. *Educational Psychologist*, *24*(2), 113–142.

Simon, S., & Maloney, J. (2007). Activities for promoting small group discussion and argumentation. *School Science Review*, *88*(324), 49–57.

Simonneaux, L. (2001). Role-play or debate to promote students' argumentation and justification on an issue in animal transgenesis. *International Journal of Science Education*, *23*(9), 903–927. doi:10.1080/09500690010016076

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory*, *6*, 174–215. doi:10.1037/0278-7393.6.2.174

Spitzhoff, D., Ramirez, S., & Wills, T. A. (1982). *The Decision Skills Curriculum: A program for primary prevention of substance abuse*. New York, NY: American Health Foundation.

Stanovich, K. E., West, R. F., & Toplak, M. E. (2012). Judgment and decision making in adolescence: Separating intelligence from rationality. In V. F. Reyna, S. B. Chapman, M. R. Dougherty, & J. Confrey (Eds.), *The adolescent brain: Learning, reasoning, and decision making* (pp. 337–378). Washington, DC: American Psychological Assocation.

Stein, M., Carnine, D., & Dixon, R. (1998). Direct instruction integrating curriculum design and effective teaching practice. *Intervention in School and Clinic*, *33*(4), 227–233.

Teasley, S. D. (1995). Communication and collaboration: The role of talk in children's peer interactions. *Developmental Psychology*, *31*(2), 207–220.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*, 453–458. doi:10.1126/science.7455683

Von Winterfeldt, D., & Edwards, W. (1986). *Decision analysis and behavioral research* (Vol. 1). Cambridge, UK: Cambridge University Press.

Vygotsky, L. S. (1962). *Thought and language*. Cambridge, MA: M.I.T. Press.

Webb, N. M., & Favier, S. (1999). Developing productive group interaction in middle school mathematics. In A. M. O'Donnell & A. King (Eds.), *Cognitive perspectives on peer learning* (pp. 117–149). Hillsdale, NJ: Erlbaum.

Weber, E. U., & Johnson, E. J. (2009). Mindful judgment and decision making. *Annual Review of Psychology*, *60*, 53–85. doi:10.1146/annurev.psych.60.110707.163633

Wegerif, R., Mercer, N., & Dawes, L. (1999). From social interaction to individual reasoning: An empirical investigation of a possible socio-cultural model of cognitive development. *Learning and Instruction*, *9*(6), 493–516. doi:10.1016/S0959-4752(99)00013-4

Wells, G., & Arauz, R. M. (2006). Dialogue in the classroom. *Journal of the Learning Sciences*, *15*(3), 379–428.

Whitehead, A. N. (1957). *The aims of education*. New York, NY: Macmillan. (Original work published 1929)

Wolfe, C. R., & Britt, M. A. (2008). The locus of the myside bias in written argumentation. *Thinking & Reasoning*, *14*, 1–27, doi:10.1080/13546780701527674

Zhang, X., Anderson, R. C., Dong, T., Nguyen-Jahiel, K., Li, Y., Lin, T., & Miller, B. (2013). Children's moral reasoning: Influence of culture and collaborative discussion. *Journal of Cognition and Culture*, *13*, 497–516. doi:10.1163/15685373-12342106

Manuscript received September 14, 2014
Final revision received July 13, 2015
Accepted August 14, 2015