Reading and Writing (2019) 32:2223–2249
https://doi.org/10.1007/s11145-019-09946-7

Effects of collaborative small-group discussions on early adolescents' social reasoning

Tzu-Jung Lin[1] · Seung Yon Ha[1] · Wei-Ting Li[2] · Ying-Ju Chiu[3] · Yu-Ru Hong[2] · Chin-Chung Tsai[4]

Published online: 27 March 2019

Abstract

This study examined the influence of small-group discussions on early adolescents' social reasoning development. A total of 147 fifth-grade students (79 males and 68 females) from six classrooms in a public school in Taiwan participated in a pre-post control quasi-experimental study. Classrooms of students were assigned to either a 5-week collaborative social reasoning (CSR) condition or an active-control read-aloud (RA) condition. All students completed a social reasoning essay before and after the intervention. Students in the CSR condition generated more social knowledge, considered more possible solutions to the complex social-moral issue, and reflected on more cognitive perspectives of the story characters. Students in the RA condition generated more shallow interpretations and were more attuned to affective perspectives of the story characters. CSR students' social reasoning tended to be more coherent, complex, and involve knowledge coordination. These findings lend support to the claim that CSR discussion is a productive vehicle for enhancing students' social reasoning.

Keywords Collaborative small-group discussion · Dialogic inquiry · Social reasoning

✉ Tzu-Jung Lin
lin.1653@osu.edu

[1] Department of Educational Studies, Ohio State University, 145A Ramseyer Hall, 29 W. Woodruff Ave., Columbus, OH 43210, USA

[2] Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taipei, Taiwan

[3] Graduate Institute of Applied Science and Technology, National Taiwan University of Science and Technology, Taipei, Taiwan

[4] Program of Learning Sciences and Institute for Research Excellence in Learning Sciences, National Taiwan Normal University, Taipei, Taiwan

Introduction

Social reasoning, generally defined as a social-cognitive capacity to consider, interpret, and weigh upon multiple aspects of a complex social situation, is essential to individuals' ability to account for their own social decisions as thoughtful citizens (Mulvey, 2016). Although decades of research has suggested that engaging in dialogue about ethics, morality, or social conventions can promote students' social or moral reasoning (e.g., Berkowitz, Gibbs, & Broughton, 1980; Damon & Killen, 1982; Mischo, 2005; Zhang et al., 2013), educational initiatives to date can at best demonstrate a weak contribution of social-moral discussions to students' social and moral development (Cheung & Lee, 2010). How to effectively incorporate social-moral discussions in classroom instructions to enhance students' social-cognitive capacity remains an open question (Nucci, Creane, & Powers, 2015).

The current study exerted a renewed effort to improve early adolescents' social reasoning through a collaborative small-group discussion approach employed in upper elementary English Language Arts classrooms. The central hypothesis of this study is that dialogic inquiry featuring argumentation and social perspective taking according to an open, collaborative, and equitable social norm can enhance students' social reasoning. This study focused on early adolescents (10–11 years) because this is a developmental period during which students begin to experience greater changes in their peer relationships, a need for social inclusion and social identities (Oberle, Schonert-Reichl, & Thomson, 2010). Given the rapid change in their social experiences, early adolescents' social reasoning may be more malleable during this stage of development. The proposed dialogic inquiry intervention may serve as a catalyst for this change.

Theories of social reasoning

Human interactions are omnipresent in our daily life and involve complex social dilemmas, such as choosing whom to befriend, which social groups to be a part of, or whether to assist a person in need. To deal with complex social situations, people rely on social reasoning to disentangle intricate relationships between their own and others' internal states, social actions, and the social contexts in which they live.

In this study, we conceptualized social reasoning as three interrelated but distinct dimensions: social knowledge, social information processing, and social perspective taking. Individuals count on social knowledge—organized schema of knowledge involving human understanding of morality, social relationships, and their surrounding environments—to handle social-moral issues in the world. Moral development theories, based on the work of Piaget (1932) and later advanced by Kohlberg (1973) and Kohlberg and Hersh (1977), suggest that children's social knowledge develops as they become able to distinguish morality from social conventions and can mentally coordinate these different perspectives (Lourenço, 2014). Social domain theorists (Turiel, 1983; Smetana, 2006) later proposed that children as young as five can conceptually distinguish between three types of social knowledge—*societal, moral,* and *personal*. Societal knowledge involves knowledge of social authority, tradition,

norms, and expectations. Moral knowledge involves the understanding of moral principles and virtues. Personal knowledge pertains to knowledge about individually oriented needs/circumstances that usually do not involve the perspectives of others. Later, some researchers (e.g., Chen-Gaddini, 2012; Dahl & Kim, 2014; Tisak, 1993) included an additional domain of social knowledge called *pragmatic consideration*, which involves practical considerations such as convenience.

While the moral development theories depict what constitutes a person's *knowledge* about a social situation, social information processing theory (Crick & Dodge, 1994) outlines a real-time *process* by which people apply social knowledge to evaluate and respond to complex social cues. This process involves several information processing strategies, ranging from attending to and encoding social cues, interpreting social cues using various reasoning strategies (e.g., analogical reasoning, logical inferencing), clarifying and selecting social goals for the social situation, generating and evaluating possible solutions to the situation, to enacting the selected solutions. Research suggests that limited social information processing skills (e.g., failure to attend to social cues, misinterpretation of social cues, aggressive response generation or response evaluation) can lead to misconduct and peer conflict (Fraser et al., 2005; Lansford, Malone, Dodge, Pettit, & Bates, 2010).

Another dimension of social reasoning pertains to *social perspective taking*. Johnson (1975) defined it as "the ability to understand how a situation appears to another person and how that person is reacting cognitively and emotionally to the situation. It is the ability to put oneself in the place of others and recognize that other individuals may have points of view different from one's own" (p. 241). Social perspective taking is crucial for the development of moral reasoning (Myyrya, Juujärvi, & Pesso, 2010), social collaboration (Walker, Shore, & Tabatabai, 2013), emotional regulation (Bengtsson & Arvidsson, 2011), and social relationships (Smith & Rose, 2011). Although the definitions and measurements vary across studies (Diazgranados, Selman, & Dionne, 2016), for the purpose of this study, we considered social perspective taking as the ability to differentiate one's own from others' cognitive or affective mental states (e.g., Sutton, Smith, & Swettenham, 1999). Although Crick and Dodge (1999) argued that social perspective taking could be one possible stage of social information processing, there has not been a concrete way to incorporate social perspective taking in the framework of social information processing. Hence, in the current study it was considered as a separate construct of social information processing.

In sum, while social knowledge reflects the *content* or *schema* of social reasoning, social information processing and social perspective taking represent the *reasoning process* and *strategies* of social reasoning. These processes and strategies of social reasoning are operated based on individuals' social knowledge.

Dialogic inquiry and social reasoning

Dialogic inquiry is a process by which students and the teacher collaboratively engage in reflective knowledge construction through argumentative talk (Alexander, 2006; Mercer, Wegerif, & Dawes, 1999; Wells, 1999). According to Walton (1998),

dialogue that takes place in the form of *inquiry* carries the goal to collaboratively develop the most reasonable solution to a controversial issue based on justifiable reasons and evidence. During dialogic inquiry, students can freely discuss open-ended questions upon which predetermined answers are not imposed. With minimal authoritative control and optimal student autonomy and support from peers and the teacher, students learn to defend their own positions and reasonably criticize others' arguments. In such an intellectually collaborative environment, multiple perspectives are welcomed, personal biases are examined, various reasoning strategies are practiced, explicated, and reflected through talk, and a socially shared understanding is developed (Mercer & Dawes, 2008; Nussbaum & Schraw, 2007; Reznitskaya et al., 2009).

We propose that dialogic inquiry about complex social-moral issues can improve students' social reasoning for the following reasons. First, the epistemic goal of dialogic inquiry—to pursue the most justifiable answer to the social-moral issues—can motivate students to understand different perspectives and make efforts to resolve conflicting perspectives among peers. Second, through argumentative talk, students first learn to socially negotiate multiple perspectives with others, and then internally coordinate the knowledge into a coherent social knowledge schema, which is the most challenging aspect of social reasoning (Kohlberg, 1973; Kohlberg & Hersh, 1977; Piaget, 1932; Smetana, 2006; Turiel, 2006). Third, during dialogic inquiry, students have opportunities to observe their peers' social information processing or social perspective taking strategies and enact them to reinforce or modify their social understanding. Fourth, the open space that is created through dialogic inquiry can allow multiple perspectives to be heard and critically examined. As students are exposed to different perspectives that they would not otherwise consider, they collectively gain conscious awareness of each others' feelings, goals, intentions, dispositions, or thought processes, thus promoting their social perspective taking.

Collaborative social reasoning discussion

Collaborative Social Reasoning (CSR) discussion is a small-group dialogic inquiry approach informed by the substantial literature on Collaborative Reasoning (CR) (Chinn, Anderson, & Waggoner, 2001; Reznitskaya et al., 2009; Sun, Anderson, Lin, & Morris, 2015), a form of teacher-scaffolded, peer-led, small-group, dialogic instruction that has revealed its significant influence on students' cognitive (Lin et al., 2015; Ma et al., 2017; Morris et al., 2018; Reznitskaya et al., 2009) and social (Sun, Anderson, Perry, & Lin, 2017) development. CSR adapts several dialogic inquiry features from CR. The first principle, as described above, is *collaborative argumentation,* whereby students rationally evaluate multiple perspectives of an issue with the goal of identifying a more defensible argument, rather than persuasive argumentation of which the goal is to compete for the most persuasive but not always the most reasonable argument. Students are prompted by the teacher to consider multiple perspectives from their group members and story characters.

Building upon recent CR findings that positive social experiences with peers contribute to the productivity of a dialogic discussion (Lin et al., 2015), the

second principle of CSR pertains to a set of social norms featuring *open discussion, equal participation,* and *mutual respect*. Through open discussion, students can freely exchange ideas with each other without constant scrutiny from the teacher (Howe, 1990). They speak without raising their hands or being called on by the teacher. Compared to traditional teacher-led direct instruction, it has been suggested that open discussion fosters a greater sense of agency and a fuller concept of audience as students construct their reasoning and arguments (Morris et al., 2018). CSR positions students on an equal footing at the center of the activity such that they all have opportunities to contribute to the discussion (Cohen & Lotan, 1995). With this greater sense of agency, students learn to regulate their talk and invite others to talk in a friendly manner to ensure that their discussion is not dominated by a few members. When students generate disagreements, they are encouraged to disagree with their opinions and not with the person, thus showing mutual respect to one another. These design principles aim at creating a collaborative, reciprocal, and supportive social space that is conducive to dialogic inquiry.

CSR extends the scope of CR to the field of social cognition by situating students in a discourse via both the abovementioned critical-analytic stance and an expressive stance—making emotive connections to the experience with the text (Murphy, Wilkinson, Soter, Hennessey, & Alexander, 2009). Students are encouraged to practice social perspective taking by sharing personal experiences or feeling relevant to the stories with peers, comparing others' experiences or feelings with their own, projecting multiple experiences or feelings from the story to make sense of story characters' intentions, feelings, or social behavior. Previous research (e.g., McVee, 2014) suggests that students can practice their social perspective taking by making such intertextual comparisons and self-other positioning.

The CSR group discussions in this study were specifically structured around issues of social exclusion and injustice. These issues are multifaceted, requiring students to consider moral principles, personal concerns, social-conventional conflicts, and practical limitations. The stories used to introduce these issues served as a medium to provoke students' social reasoning within a safe, literary environment.

In this study, the effects of CSR on students' social reasoning development was examined by comparing the post-intervention reflective essays of the students in the CSR group and those in an active-control Read-Aloud group, controlling for their pre-intervention social reasoning, academic performance, and gender. The major difference between the CSR and Read-Aloud groups lies in the process of dialogic inquiry. Both groups read the same story texts throughout the intervention. While the CSR group participated in the dialogic inquiry discussions described above, the Read-Aloud group mainly experienced traditional teacher read-aloud practices and individual writing activities. Students' social reasoning was assessed in the form of a reflective essay, with an underlying assumption that reasoning skills can be transferred from oral discussion to individual writing (Reznitskaya et al., 2001). It has been suggested that essay writing provides more open space for thoughts and fewer prompts from experimenters, which enables a better understanding of students' spontaneous reasoning processes (e.g., Nucci et al., 2015; Zhang et al., 2013) compared to questionnaires or interviews (e.g., Damon, 1980; Park & Killen, 2010).

Two research questions were addressed in this study: (1) Does the CSR group generate more social knowledge, social information processing strategies, and social perspective taking than the Read-Aloud group in their post-intervention essays? (2) Does the CSR group present their social reasoning in writing with greater coherence, complexity, and knowledge coordination than the Read-Aloud group? We hypothesized that the CSR discussions would foster greater development of social reasoning than the Read-Aloud instruction. Specifically, we expected that the CSR students would generate more social knowledge using various social information processing and social perspective taking strategies than the Read-Aloud students; the ideas in the essays would be more meaningfully connected (i.e., coherence), demonstrating multiple perspectives (i.e., greater complexity) in a coordinated manner (i.e., knowledge coordination).

Method

Participants

This study included 147 fifth-grade students (79 males and 68 females, $M_{age} = 11.10$, $SD_{age} = 0.31$) from six classrooms in an urban public elementary school serving students from middle-class families in Taiwan. Teachers reported that three of the students had attention deficit hyperactivity disorder (ADHD), eight had a learning disability, and two had Asperger syndrome (AS). All of the teachers (5 female, 1 male) in the six classrooms were homeroom teachers. Their teaching experience ranged from 5 to 19 years ($M = 11.83$).

Teacher training

The CSR teachers attended a whole-day workshop prior to the intervention, received weekly coaching and consultation throughout the 5-week intervention, and participated in a mid-intervention teacher meeting with researchers. During the whole-day workshop, the teachers learned about theories and evidence supporting the design features of CSR, as well as strategies to set up a discussion and facilitate social reasoning and participation during the discussion. Throughout the intervention, field researchers audio- and video-recorded all of the discussions and provided on-site suggestions, such as the seating arrangements. After each week's discussion, the researchers reviewed the videos as a team and provided each teacher with written feedback via email or instant messages. After the second week of intervention, a mid-intervention teacher meeting was held to discuss major instructional issues that the teachers encountered and potential instructional strategies that they could use to meet the students' needs. The Read-Aloud teachers received a half-day training prior to the intervention to learn about the implementation procedure and story materials. These teachers showed the same interest as the CSR teachers in implementing CSR in their classrooms, and received all CSR training and materials after the study.

Study condition and procedure

This study followed a pre-post control quasi-experimental design. Three of the classrooms were assigned to the CSR condition and the other three were assigned to the Read-Aloud Condition. Classroom assignment was based on the criterion that students from the two conditions had matched fourth-grade academic achievement. All of the students read one story featuring complex social exclusion or injustice issues each week for five consecutive weeks. The story order was counterbalanced across classrooms based on a semi-Latin-Square design.

Collaborative social reasoning condition (CSR)

Each class was divided into four heterogeneous small groups, each a cross-section of the classroom in terms of gender, ethnicity, achievement, social skills, and social relationships. One or two weeks prior to the small-group intervention, each teacher held a CSR set-up session in their classroom in which students were shown a 5-min video of a group of students discussing a school policy issue. After the video, the teachers and students engaged in a whole-class discussion about the video, centering on how the students in the video conformed to CSR's norm and ground rules. Each classroom then set up their own ground rules for the upcoming CSR discussions based on the principles.

Each week, before the discussion, the CSR students read a story independently or with a classmate. After reading, a goal-setting activity was implemented to foster the students' motivation and engagement in the dialogic inquiry. Previous research suggested that the types of goals students or the teacher set for the discussion can influence their ways of thinking and participation during the discussion (Nussbaum, Kardash, & Graham, 2005). Students were asked to fill out a goal-setting sheet to set goals for argumentation and participation for the upcoming discussion. Examples of students' self-generated goals are "Respect each other," "Look at multiple perspectives" and "Be a better listener."

A CSR discussion began with a teacher-led introduction, in which teachers invited students to share their individual goals and involved all group members to develop their group goals based on the ground rules that the whole class set up at the beginning of the semester. After goal setting, the teachers announced the big question—a central social-moral dilemma raised from the story—to initiate the discussion. Students shared their initial positions on the big question, followed by a free-flowing argumentative discussion. The teacher acted as a facilitator, scaffolding students' argumentation and social participation when necessary. Teachers were provided with a list of supplemental questions (one question per story) during the teacher workshop. The purpose of these supplemental questions was to help students connect the story to their personal experiences. Teachers could adapt these questions freely when using them during a discussion.

On average, each CSR discussion lasted for 21.73 min. The body of the discussion ended with students' individual final positions, followed by a teacher-led debriefing session where students reflected on their individual and group

performance in respect to their self-generated goals. The group then set goals for the next discussion.

Read-aloud condition (RA)

Each week, the teacher read one story to their students in a whole-class setting. The stories used in the CSR and RA conditions were identical. Students reflected on the big question and its supplemental question individually in writing at their own desk. On average, the writing task took 30 min.

Literature-based social reasoning unit

A 5-week literature-based social reasoning unit was developed for the CSR and RA students in this study. Five fictitious stories at the fifth-grade reading level were selected for the study. The issues raised from the stories were provocative and related to social exclusion and injustice. These stories were excerpted or adapted from Diary of a Wimpy Kid (Kinney, 2013), There is a Boy in the Girls' Bathroom (Sachar, 1987), Harriet the Spy (Fitzhugh, 2009), Wonder (Palacio, 2012), and Bat 6 (Wolff, 2000). For example, the story Bat 6 describes two teams of sixth-grade girls who are ready for their annual Bat 6 softball game in the context of the United States after World War II. This year each team has a newcomer. Aki has just returned with her family from an American Internment Camp. Shazam has been shunted around by her mother since her father died at Pearl Harbor. On the day of Bat 6, Shazam attacks Aki and breaks her jaw. The big question was "Should Aki forgive Shazam?," and the supplemental question was "How should we treat people with different cultural backgrounds?"

Measure

Social reasoning essay task

Students received a social reasoning essay task before and after the intervention to reason about a social exclusion issue arising from a short story. The research team developed two stories and randomly distributed them to students as the pre- and post-test. The order of the story was counterbalanced among students within classrooms. One of the stories, Private! (723 words) was adapted from an online resource: Daily Dilemma: *The Situation* by Denison (2014). This story describes a new student, David, who finds that his group of new friends has uploaded mocking photos of a classmate in a private online group. The writing prompt for this story was "Should David report the post to the school principal?" The second story, Super-Sized Slugger (746 words) was adapted from a novel of the same title (Ripken & Cowherd, 2012). This story describes a boy, Cody, who is being bullied by a member of his baseball team, Dante, for his physical appearance. The writing prompt for this story was "Should Cody tell on Dante?"

Fourth-grade academic achievement

Students' fourth-grade academic achievement scores averaged across all major subject areas (Language Arts, Math, Science, Social Studies) were obtained from the school teachers.

Essay coding and analysis

Students' social reasoning essays were digitized and presented to the raters with numeric identifiers that concealed the study condition information. To evaluate students' social reasoning, we developed (1) an analytic coding scheme to examine students' social knowledge, social information processing strategies, and social perspective taking strategies, and (2) a holistic scoring rubric to assess the essays' coherence, complexity, and knowledge coordination.

Analytic coding

Each essay was segmented into communication units (c-units) (Loban, 1976), defined as "an independent clause with its modifiers" (p. 9), which, in the current study, represented the smallest unit of a complete and independent idea. Three trained researchers independently segmented all of the pre-test and post-test essays based on a segmentation convention developed by the research team. Approximately 20% of the randomly selected essays were segmented by the three researchers for reliability checking. The inter-rater agreement among each pair of the three researchers ranged from 82.8% to 96.0%. All segmented essays were then imported into the QSR International's NVivo11 software (QSR, 2015) for coding.

The multidimensional social reasoning coding scheme (Table 1) included three dimensions: social knowledge, social information processing, and social perspective taking. Each c-unit generated in the essay could receive at most one code from each dimension. Each dimension consisted of macro-level categories informed by the literature (see below for more explanations). We used a bottom-up approach to identify specific topics or strategies that characterized each macro-level category, hereafter called micro-level categories, based on the researchers' examination and discussion of 10% of the randomly selected essays. Some of the macro-level categories did not contain micro-level categories because they could not be reliably identified. Essays were coded by two independent researchers. The inter-rater reliability across all categories based on 20% of the essays was satisfactory (mean Cohen's Kappa = 0.89, range = 0.65–1.00, see Table 1 for more details).

Based on the moral development theories (Kohlberg, 1973; Kohlberg & Hersh, 1977; Piaget, 1932; Smetana, 2006; Turiel, 2006), written social knowledge was coded into one of four types of social knowledge: societal knowledge, moral knowledge, personal knowledge, and pragmatic consideration. Societal knowledge involved topics related to authority, tradition, norms, and expectations established by the society (e.g., culture, school, peer group). Personal knowledge involved

Table 1 Analytic social reasoning coding scheme

Dimension	Macro-level category	Micro-level category	Examples of coded responses	Cohen's Kappa
Social knowledge	Societal knowledge	Friendship fidelity	They are his best friends, so he shouldn't rat them out completely	0.92
		Group advantage	The team is more likely to win the game if he stays on the team	
		Group disadvantage	Telling on Dante might lead the team to lose the game	
		Peer pressure	If she stands up then they won't like her anymore	
		Politeness or rudeness	Asking someone to sit in another place is rude	
		Punishment	His friends might get expelled if he reports them	
		Rules or laws	Cyber-bullying is against school policies	
		Appeal to authority	Cody should talk to his parents first	
		Social or cultural norms	Classmates are supposed to be friendly to each other	
	Personal knowledge	Personal advantage	If he tells the principal, he can get even more friends	0.81
		Personal disadvantage	If Cody does not tell on Dante, he may be called a coward forever	
		Personal concern	Dante might be aggressive because he is afraid of losing his position	
		Personal choice	Everyone has his/her own thoughts	
	Moral knowledge	Basic human rights	Nobody should be judged on their appearance	0.86
		Fairness	Everybody should be given the same opportunity	
		Forgiveness	Brandon should be given a second chance	
		Honesty	Telling a lie is wrong	
		Moral responsibility	When someone is being bullied, you just have to help them out so that this won't happen again	
		Moral judgment	Making fun of someone is not the right thing to do	
	Pragmatic consideration		I also think he should delete Brandon from his Facebook and delete his Facebook entirely	0.65

Table 1 (continued)

Dimension	Macro-level category	Micro-level category	Examples of coded responses	Cohen's Kappa
Social information processing	Encoding		The story says that David was a new kid in the school	1.00
	Interpretation	Alternative hypothesis	Cody only thought of keeping the position to himself and did not think of what they could earn if they collaborated	1.00
		Judgment	Bullying is bad, and bullies will get punished	
		Logical inference	The worry for Cody was that Dante had been the third-baseman last year, so he might not have a chance to play	
		Analogy		
		(Surface Analogy)	Speaking of the Internet, people like to gossip on the Internet. [associating the term 'Internet' from the text with the student's knowledge of how the Internet impacts our life without linking this knowledge to the student's main argument]	
		(Surface + Relational Analogy)	Bullying is not tolerable, because I was once bullied and that hurt. [associating the bullying scenario in the text with students' own bullying experience]	
		(Relational Analogy)	This feeling is like you want a job so eagerly but then someone shows up to compete with you. [mapping the stressful feeling in the text with the stressful feeling that one may have when having to compete for a dream job]	
	Response generation		I don't think David should report the post to the school principal Cody should tell on Dante	1.00
	Response evaluation		Even if it means losing all his friends, it's still the right thing to do	1.00
Social perspective taking	Affect perspective	Understanding of emotion or feeling	Who would not feel sad if they were Samantha?	0.75
	Cognitive perspective	Understanding of goals, intention, dispositions, thought processes	People might bully someone because they want to have what he has	0.75

topics such as individuals' internal thoughts, motivation, feeling, or concerns. Moral knowledge involved topics related to the judgment of universal moral principles. Pragmatic consideration referred to the utility or consequences of an action. Based on social information processing theory (Crick & Dodge, 1994), students' social reasoning was broken down into four social information processing strategies: encoding, interpretation, response generation, and response evaluation. Although behavioral enactment was considered the last step of social information processing, we did not include this process in our coding scheme because it was less possible for students to present this process in a written essay. Encoding referred to the perception of and attention to social cues, which could be external (e.g., situational and contextual) or internal (e.g., emotional and personal). C-units that were coded as encoding were those that showed students' comprehension of textual information or emotional arousals (e.g., "I feel bad about David"). Interpretation referred to reasoning strategies by which meanings were derived from the social cues. The interpretation category was further coded into several reasoning strategies drawn from previous CR studies of students' moral and relational reasoning (Lin et al., 2012, 2015; Zhang et al., 2013; Zhang et al., 2016): *Alternative hypothesis* referred to generating new hypotheses or interpretations of the presented social-moral issue other than the original ideas presented in the students' essay or in the story. *Judgment* referred to social-moral evaluation of someone's social conduct, which was often reflected in the students' essays as judgment of whether a particular conduct or attitude was right or wrong. *Logical inference* referred to interpretation involving causal relationships (e.g., premise and consequence) among the social cues. *Analogy* referred to mapping the similarity between the information from the text and from their prior knowledge. Analogies were classified into *surface analogy* when students related fragmented information from the text to their prior knowledge. *Surface + relation analogy* referred to the similarity alignment between a coherent piece of information from the text and their prior knowledge. *Relational analogy* referred to a thematic alignment between a knowledge-based argument and the student's prior knowledge. Response generation referred to the process by which desired goals were formulated in action based on the student's interpretation of the situation. Lastly, potential responses were evaluated based on the student's capability to respond and the consequences of the decision, called response evaluation.

Social perspective taking was separated into affective perspective taking and cognitive perspective taking. *Cognitive perspective taking* referred to understanding others' goals, intentions, dispositions, and thought processes. *Affective perspective taking* referred to an understanding of others' emotions or feelings.

Holistic scoring

A holistic scoring rubric was informed by two fields of research: (1) moral development theories (Kohlberg, 1973; Kohlberg & Hersh, 1977; Piaget, 1932; Smetana, 2006; Turiel, 2006), which suggest that advanced social reasoning must involve a complex process of knowledge coordination; (2) collaborative argumentation theories (Reznitskaya et al., 2009; Golanics & Nussbaum, 2008), which suggest that a good reasoner is able to generate reasonable arguments by drawing logical

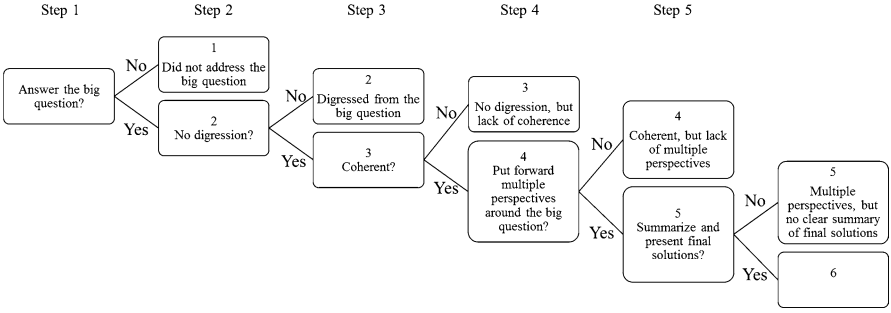

Fig. 1 A five-step decision tree for the holistic scoring of social reasoning essays

connections between claims, evidence, reasons, counterarguments, and rebuttals. The rubric was set up in the form of a five-step decision tree, with three particular focuses—coherence, complexity, and knowledge coordination (Fig. 1). In Step 1, the raters evaluated whether the students attempted to address the *big questions* for the essay task (similar to the big questions provided during the intervention). If the attempt was identified, in Step 2, the raters evaluated whether any idea generated digressed from the big question. If all ideas were relevant to the big question, in Step 3, the raters evaluated whether the student's ideas were locally coherent. An argument was locally coherent when each of the main ideas was supported or elaborated by logical reasons and evidence. If local coherence was reached, in Step 4, the raters evaluated whether the student's ideas involved multiple perspectives, an index of complexity, mainly in the forms of counterargument and rebuttals. We considered topic relevance and local coherence as the prerequisites of multiple perspectives in order to distinguish essays including multiple fragmented ideas from essays including multiple reasonable ideas. If the essay was considered complex, in Step 5, the raters evaluated whether the student concluded his/her ideas with a justifiable solution to the big question, an index of knowledge coordination. The two raters were trained in two rounds to obtain a satisfactory inter-rater reliability. In the first round, the raters independently scored 10% of the essays. The Kappa value was 0.87. All of the disagreements were resolved. In the second round, the raters independently scored another 10% of the essays and achieved a Kappa value of 0.90.

Results

Preliminary analysis

Table 2 presents the means and standard deviations of both analytic and holistic social reasoning measures for the pre- and post-tests. A two-sample *t* test was used to test if the CSR and RA groups were equivalent at the baseline. The results showed that the two groups did not differ in their fourth-grade academic achievement, supporting the baseline equivalence assumption. However, a Poisson regression analysis comparing the total number of c-units in the pre-test essays of the CSR and RA groups suggested that the CSR group generated more c-units than the

Table 2 Means and standard deviations of students' social reasoning

Social reasoning variables	CSR ($n=71$)		RA ($n=76$)	
	Pretest	Posttest	Pretest	Posttest
	M (SD)	*M* (SD)	*M* (SD)	*M* (SD)
Analytic coding				
Social knowledge	5.13 (2.51)	8.24 (5.24)	3.20 (2.42)	6.18 (3.22)
Societal	0.75 (0.85)	1.35 (1.45)	0.74 (1.09)	0.75 (1.05)
Personal	1.69 (1.87)	2.82 (2.56)	1.03 (1.08)	2.37 (2.52)
Moral	2.23 (1.70)	3.08 (2.75)	1.24 (1.33)	2.46 (1.75)
Pragmatic consideration	0.46 (0.94)	0.99 (1.49)	0.20 (0.46)	0.61 (1.02)
Social information processing	8.14 (3.72)	12.55 (6.84)	6.16 (3.71)	11.61 (4.58)
Encoding	1.62 (1.99)	2.32 (2.47)	1.14 (2.00)	2.00 (2.42)
Interpretation	3.08 (2.48)	4.15 (4.06)	1.91 (2.19)	4.28 (3.13)
Alternative hypothesis	0.08 (0.33)	0.10 (0.42)	0.04 (0.20)	0.16 (0.67)
Judgement	1.93 (1.42)	2.32 (2.46)	1.22 (1.62)	1.83 (1.76)
Logical inference	0.70 (1.10)	1.20 (1.69)	0.32 (0.57)	1.17 (1.46)
Analogy	0.37 (1.12)	0.54 (2.05)	0.33 (1.23)	1.12 (2.77)
Surface analogy	0.18 (0.95)	0.15 (1.09)	0.16 (0.97)	0.67 (2.52)
Surface + relational analogy	0.18 (0.66)	0.37 (1.42)	0.11 (0.56)	0.38 (1.34)
Relational analogy	0.00 (0.00)	0.01 (0.12)	0.07 (0.57)	0.05 (0.36)
Response generation	1.99 (1.15)	3.68 (2.67)	1.80 (1.06)	2.79 (1.59)
Response evaluation	1.45 (2.07)	2.39 (2.81)	1.30 (1.47)	2.54 (3.30)
Social perspective taking	1.42 (2.60)	2.48 (3.09)	0.83 (1.99)	3.22 (5.17)
Affective perspective	0.65 (1.52)	0.73 (2.01)	0.43 (1.10)	1.87 (4.76)
Cognitive perspective	0.77 (2.04)	1.75 (2.12)	0.39 (1.49)	1.36 (2.34)
Holistic scoring	3.17 (1.24)	3.99 (1.37)	3.04 (1.25)	3.34 (1.19)

*$p<0.5$; **$p<0.01$; ***$p<0.001$

RA group ($M_{CSR}=8.52$, $SD_{CSR}=4.00$, $M_{RA}=6.13$, $SD_{RA}=3.66$, $B=0.33$, $SE=0.06$, Wald $\chi^2=28.51$, $p<0.001$). In particular, the CSR students generated more social knowledge ($B=0.45$, $SE=0.08$, Wald $\chi^2=30.04$, $p<0.001$), social information processing ($B=0.28$, $SE=0.06$, Wald $\chi^2=20.15$, $p<0.001$), and social perspective taking ($B=0.53$, $SE=0.18$, Wald $\chi^2=8.55$, $p<0.01$) in the pre-test than the RA group. These pre-test differences might be a result of the CSR students' heightened motivation and excitement after the researchers introduced the CSR activities in a student recruitment and consent session. Pre-test social reasoning performance was included in later analyses to control for the pre-intervention difference.

Social knowledge, social information processing, and social perspective taking

To examine whether the CSR group generated more social knowledge, social information processing, or social perspective taking than the RA group, Poisson

regression analyses were conducted in which post-test social reasoning was predicted by condition (CSR vs. RA), with the control of the pre-test scores, fourth-grade academic achievement, and gender.

As shown in Table 3, the total amount of social knowledge was greater in the CSR group than in the RA group (least squared $M_{CSR}=2.01$, $SE_{CSR}=0.05$; least squared $M_{RA}=1.87$, $SE_{RA}=0.05$; $p<0.05$), particularly for societal knowledge (least squared $M_{CSR}=0.20$, $SE_{CSR}=0.11$; least squared $M_{RA}=-0.29$, $SE_{RA}=0.13$; $p<0.01$) and pragmatic consideration (least squared $M_{CSR}=-0.09$, $SE_{CSR}=0.12$; least squared $M_{RA}=-0.50$, $SE_{RA}=0.15$; $p<0.05$). The two treatment groups did not differ in their moral knowledge or personal knowledge. Fourth-grade academic achievement was positively predictive of total amount of social knowledge ($p<0.001$), societal knowledge ($p<0.01$), moral knowledge ($p<0.001$), and pragmatic consideration ($p<0.05$). Gender had no effects on any of the social knowledge variables.

The total amount of social information processing did not differ by condition (Table 3). However, the CSR group was found to generate more responses or solutions to the big question than the RA group (least squared $M_{CSR}=1.27$, $SE_{CSR}=0.06$; least squared $M_{RA}=1.02$, $SE_{RA}=0.07$; $p<0.01$). The RA group generated more interpretations (least squared $M_{CSR}=1.28$, $SE_{CSR}=0.06$; least squared $M_{RA}=1.47$, $SE_{RA}=0.06$; $p<0.05$) than the CSR group. When different reasoning strategies for interpretation were further compared using Poisson regression analyses, results suggested that the RA group outperformed the CSR group in the amount of interpretation, mainly due to the greater number of analogies they generated (least squared $M_{CSR}=-0.75$, $SE_{CSR}=0.17$; least squared $M_{RA}=0.04$, $SE_{RA}=0.11$; $p<0.001$). Taking into account the types of analogies generated by the students, the RA group generated more surface analogies than the CSR group (least squared $M_{CSR}=-3.38$, $SE_{CSR}=4.05$; least squared $M_{RA}=-0.83$, $SE_{RA}=3.99$; $p<0.05$), while the two groups did not differ in the number of surface + relation analogies or relational analogies. They also did not differ in the amounts of encoding or response evaluation generated in their essays. Fourth-grade academic achievement was positively related to the total amount of social information processing ($p<0.001$), interpretation ($p<0.001$), and response generation ($p<0.05$). Girls were more likely to encode social cues from the text than boys ($p<0.01$), whereas boys were more likely to evaluate their responses than girls ($p<0.05$).

In terms of social perspective taking, after controlling for the pre-test scores, the CSR group considered more cognitive perspectives than the RA group (least square $M_{CSR}=0.35$, $SE_{CSR}=0.10$; least square $M_{RA}=0.04$, $SE_{RA}=0.11$; $p<0.05$). In contrast, the RA group considered more affective perspectives than the CSR group (least square $M_{CSR}=-0.46$, $SE_{CSR}=0.15$; least square $M_{RA}=0.25$, $SE_{RA}=0.10$; $p<0.001$).

To further understand the quantitative differences between the two groups of students' social reasoning, we selected one representative post-intervention social reasoning essay from each treatment group for a qualitative comparison. The first essay was written by a student from the CSR group. Throughout the essay, the student illustrated in detail how he would deal with the social situation if he were the main story character, David. Rather than answering the big question (Should David tell on Dante?) with a simple yes or no response, his responses carefully considered various conditions in which the event would have occurred (e.g., whether Brandon was

Table 3 Poisson regression models of students' social reasoning predicted by study condition

Outcome variable	Predictor	B	S.E.	95% Wald CI	Wald χ^2
Social knowledge					
Total social knowledge	Condition (CSR = 1, RA = 0)	0.14*	0.07	0.01 to 0.27	4.49
	Pre-test	0.05***	0.01	0.03 to 0.08	15.71
	Fourth grade Achievement	0.04***	0.01	0.02 to 0.06	21.35
	Gender (1 = Male, 0 = Female)	0.10	0.07	− 0.03 to 0.22	2.11
Societal knowledge	Condition (CSR = 1, RA = 0)	0.49**	0.17	0.16 to 0.83	8.42
	Pre-test	0.11	0.08	− 0.04 to 0.29	0.15
	Fourth grade Achievement	0.06**	0.02	0.02 to 0.11	7.15
	Gender (1 = Male, 0 = Female)	0.31	0.18	− 0.03 to 0.65	3.12
Personal knowledge	Condition (CSR = 1, RA = 0)	0.08	0.11	− 0.13 to 0.29	0.57
	Pre-test	0.10**	0.03	0.04 to 0.16	9.97
	Fourth grade Achievement	0.02	0.01	− 0.01 to 0.05	2.43
	Gender (1 = Male, 0 = Female)	− 0.15	0.11	− 0.36 to 0.07	1.79
Moral knowledge	Condition (CSR = 1, RA = 0)	0.13	0.10	− 0.07 to 0.34	1.57
	Pre-test	0.04	0.03	− 0.03 to 0.10	1.36
	Fourth grade Achievement	0.06***	0.01	0.03 to 0.09	18.40
	Gender (1 = Male, 0 = Female)	0.19	0.11	− 0.02 to 0.40	3.20
Pragmatic consideration	Condition (CSR = 1, RA = 0)	0.41*	0.20	0.03 to 0.80	4.40
	Pre-test	0.12	0.10	− 0.08 to 0.32	1.37
	Fourth grade Achievement	0.06*	0.03	0.01 to 0.11	5.04
	Gender (1 = Male, 0 = Female)	0.31	0.20	− 0.08 to 0.71	2.44
Social information processing					
Total social information processing	Condition (CSR = 1, RA = 0)	− 0.02	0.05	− 0.12 to 0.08	0.20
	Pre-test	0.03***	0.01	0.02 to 0.05	23.14
	Fourth grade Achievement	0.03***	0.01	0.02 to 0.04	18.72
	Gender (1 = Male, 0 = Female)	0.03	0.05	− 0.07 to 0.13	0.37
Encoding	Condition (CSR = 1, RA = 0)	0.14	0.12	− 0.09 to 0.36	1.44
	Pre-test	− 0.08**	0.03	− 0.15 to − 0.02	6.71
	Fourth grade Achievement	0.03	0.02	0.00 to 0.06	3.26
	Gender (1 = Male, 0 = Female)	− 0.35**	0.12	− 0.58 to − 0.11	8.10

Table 3 (continued)

Outcome variable	Predictor	B	S.E.	95% Wald CI	Wald χ^2
Interpretation	Condition (CSR = 1, RA = 0)	−0.19*	0.08	−0.36 to −0.02	5.05
	Pre-test	0.08***	0.02	0.05 to 0.11	24.57
	Fourth grade Achievement	0.05***	0.01	0.03 to 0.08	20.82
	Gender (1 = Male, 0 = Female)	0.05	0.09	−0.12 to 0.22	0.35
Response generation	Condition (CSR = 1, RA = 0)	0.24**	0.09	0.06 to 0.43	6.74
	Pre-test	0.10**	0.04	0.02 to 0.18	6.64
	Fourth grade Achievement	0.03*	0.01	0.01 to 0.05	5.45
	Gender (1 = Male, 0 = Female)	0.07	0.10	−0.12 to 0.26	0.52
Response evaluation	Condition (CSR = 1, RA = 0)	−0.11	0.11	−0.32 to 0.10	1.07
	Pre-test	0.09***	0.03	004 to 0.14	11.40
	Fourth grade Achievement	0.02	0.01	−0.01 to 0.04	1.53
	Gender (1 = Male, 0 = Female)	0.25*	0.11	0.03 to 0.48	4.93
Social perspective taking					
Total perspectives	Condition (CSR = 1, RA = 0)	−0.12	0.11	−0.33 to 0.10	1.13
	Pre-test	0.00	0.04	−0.07 to 0.07	0.00
	Fourth grade Achievement	0.05***	0.02	0.02 to 0.08	11.02
	Gender (1 = Male, 0 = Female)	−0.12	0.12	−0.35 to 0.11	1.00
Affective perspectives	Condition (CSR = 1, RA = 0)	−0.71***	0.18	−1.06 to −0.35	15.24
	Pre-test	−0.09	0.11	−0.29 to 0.12	0.66
	Fourth grade Achievement	0.00	0.02	−0.04 to 0.04	0.00
	Gender (1 = Male, 0 = Female)	−0.46**	0.19	−0.83 to −0.10	6.17
Cognitive perspectives	Condition (CSR = 1, RA = 0)	0.31*	0.15	0.02 to 0.60	4.37
	Pre-test	−0.02	0.07	−0.16 to 0.11	0.12
	Fourth grade Achievement	0.09***	0.02	0.05 to 0.14	18.44
	Gender (1 = Male, 0 = Female)	0.10	0.15	−0.20 to 0.40	0.42

the one who made the post, whether Brandon would delete the post). The multiple responses or solutions that the student generated were then carefully evaluated based on the student's social knowledge using several social information processing and social perspective taking strategies, including moral judgment, alternative hypothesis, and cognitive perspective taking.

Essay 1 from the CSR Condition [see the Chinese version in the Appendix]

> If I were David, I would first confirm who did this. [**Response Generation**] If it was not Brandon, I would notify the principal that someone posted this. [**Response Generation**] If it was Brandon I would suggest that he delete the post. [**Response Generation**] If he did not want to do that I would tell him what will happen if someone finds out. [**Response Generation/Social Perspective Taking (cognitive)**] And if he still thinks this is fun and does not want to delete the post, I will tell on him immediately. [**Response Generation/Social Perspective Taking (cognitive)**] Even though this would influence my interpersonal relationships, we should do the right thing. [**Response Evaluation/Moral Knowledge**] This has nothing to do with me; it is related to justice and we need to do things for the sake of justice. [**Interpretation (judgment)/Moral knowledge**] On the other hand, I may just ignore this post and pretend nothing has happened because I do not want to offend my good friends. [**Interpretation (alternative hypothesis)/ Societal Knowledge (peer pressure)**] Everyone says that Samantha is the teacher's pet, and no one likes her. [**Encoding/ Societal knowledge**] I don't think this is right, [**Interpretation (judgment)/Moral knowledge**] although I would keep this a secret because I am not acquainted with her and we are not friends. [**Response Generation/Personal Knowledge**]

Essay 2 was generated by a student from the RA group. Throughout the essay, the student was captivated by two main ideas: moral behavior and empathy. She argued that David should report the post because what his friends did was not moral, and they did not consider the victim's feelings. Based on these two ideas, the student commented in the end that these would lead to better friendships, although she did not provide warrants to make connections between the reasons and the conclusion.

Essay 2 from the RA Condition

> I think David should report the post to the principal. [**Response Generation**] This can prevent his classmates from distributing the post, [**Response Evaluation**] which is not a moral thing at all. [**Interpretation (judgment/Moral knowledge)**] Plus, for 'the person' who distributed the post, if other people distribute his post, I believe that he would feel very sad when he sees the post. It's easy to make others sad, but it is hard to make others happy. Even if you know someone and you see him happy from the outside, how about his inside? Have you ever thought that his heart is actually bitter and you can't imagine how painful it is. [**Interpretation (analogy)/Social Perspective Taking (affective)**] Therefore, we should stand in others' shoes. Don't focus on those bad things. As we live in this world, we should not be bystanders. [**Interpretation (judgment)/Moral knowledge**] That way we will be able to enjoy friendships. [**Response Evaluation/Personal knowledge**]

Overall, the two essays demonstrated that the RA student generated more shallow inferences from his/her own personal experiences and focused more on the story character's affect or feeling. In contrast, the CSR student generated more responses or solutions to the big question, and the interpretations were grounded in multiple perspectives. Comparatively, the CSR student's essay also demonstrated more use of social information processing strategies and cognitive perspective taking.

Coherence, complexity, and knowledge coordination

In the pre-test, 19.7% of the essays from the entire sample were rated as coherent but lacking in complexity, 3.4% of the essays contained multiple perspectives on the issue but did not include concluding statements; only 2.7% of the essays contained a summary or concluding statements after the students pondered on multiple perspectives on the issue. In the post-test, 33.3% of the essays were rated as coherent but lacking in complexity, 17.7% were coherent and contained multiple perspectives, and 11.6% contained concluding statements from multiple perspectives.

To examine whether the CSR group's social reasoning essays were more coherent, complex, or coordinated than those of the RA group, an ordinal logistics regression was conducted to predict post-test holistic essay scores by condition, controlling for pre-test holistic essay scores, fourth-grade academic achievement, and gender. Essays were rated higher in the CSR group than the RA group ($B = 0.71$, $SE = 0.31$, Wald 95% CI: [0.11, 1.31], Wald $\chi^2 = 5.38$, $p < 0.05$). The effect of the pre-test essay scores was significant ($B = 0.29$, $SE = 0.15$, Wald 95% CI: [0.00, 0.58], Wald $\chi^2 = 3.80$, $p = 0.05$). The effects of gender and academic achievement were not significant.

Below is a sample essay from a CSR student that demonstrated a coherent and complex argument. The student first used textual evidence to comment on Brandon and his friends' negative social behavior. He then explained why the behavior needed to be reported using a combination of societal, moral, and personal knowledge, as well as practical consideration. The student considered these multiple perspectives using social information processing and social perspective taking strategies. The student's claim was clearly supported by several reasons; counterarguments (i.e., afraid of losing friends) were considered and well rebutted. While considering multiple perspectives, the student evaluated several potential solutions to the issue and finally concluded with a final decision based on moral, societal, and practical reasons, demonstrating the ability to coordinate multiple perspectives into a coherent social schema.

Essay 3

> I think David should tell teachers and adults about this post, [**Response Generation**] because Brandon and other people colored Samantha's picture, teased her on social media, and humiliated her with negative words. [**Encoding/Moral Knowledge**] Samantha did nothing wrong to them. [**Interpretation (judgement)/Moral knowledge**] This means that Brandon and other people

intentionally hindered other people's freedom, just like asking others to do what he wants them to do. [**Interpretation (logical inference)/Moral knowledge**] According to the constitution and law of freedom, they have violated the law. [**Interpretation (judgment) /Societal knowledge**] Samantha has the right of participation, freedom, and privacy [**Interpretation (judgment)/ Moral knowledge**]. She can do whatever she wants to do. [**Interpretation (judgment)/Personal knowledge**] The important thing is that she did not offend others. [**Interpretation (judgment)/Societal knowledge**] Brandon and other people should have thought whether their social conduct would 'make others feel uncomfortable.' [**Response Generation/Moral Knowledge/Social Perspective Taking (affective)**] They should treat others the way they want to be treated. [**Response Generation/Moral knowledge**] When you treat others well, others will also treat you well. [**Interpretation (Relational Analogy)/Moral Knowledge**] But they did not think about it this much. They just thought that it was funny and didn't consider others' feelings. [**Interpretation (judgment)/Personal Knowledge/Social Perspective Taking (cognitive)**] So, they not only verbally bullied others, but also violated the law of freedom and privacy. [**Interpretation (judgment)/Societal knowledge**]

I think David should tell on them. [**Response Generation**] David should not set aside such an important thing just because he is afraid of losing his friends. [**Response Evaluation/Personal knowledge**] One day, when someone else reports the post, David would become an accomplice. [**Interpretation (logical inference)/Societal Knowledge**]That will do nothing good for either side. [**Response Evaluation/Practical Consideration**] It's a lose-lose situation. [**Response Evaluation/Practical Consideration**)] So, David should report the post as soon as possible. [**Response Evaluation**]

Discussion

The findings of this study support our major hypothesis that CSR discussion is a productive vehicle for enhancing students' social reasoning. Compared to the active-control RA group, students in the CSR condition generated more social knowledge, considered more possible solutions to the complex social-moral issue, and reflected on more cognitive perspectives of the story characters in their essays, after controlling for pre-test differences. Although the RA group generated more interpretations and affective perspective taking than the CSR group, their reasoning was found to be relatively shallow and intuitive. The CSR students' social reasoning essays tended to be more coherent and complex, and involved knowledge coordination. These differences in the social reasoning of the students in the CSR condition and their active counterparts can be attributed to the dialogic inquiry process during the CSR discussions, in which students socially and critically considered multiple perspectives of the issue to pursue the most reasonable solution, and practiced various social information processing and social perspective taking strategies.

The significant effect of the CSR discussions on the students' ability to consider multiple types of social knowledge have significant theoretical and practical implications for the field. Some moral development researchers have indicated that cognitive maturation limits students' ability to reason beyond their own perspectives (e.g., Gibbs, 2003; Kohlberg, 1973; Kohlberg & Hersh, 1977), while others (Nucci et al., 2015; Smetana, 2006; Turiel, 2006) have argued for children's potential to conduct complex social reasoning. These researchers have suggested that children as young as the age of three have the potential to consider and coordinate various types of social knowledge. Our findings support this latter account. Even the RA group students were able to contemplate on multiple types of social knowledge to some degree even though they were exposed to the complex social-moral issues with little social-cognitive scaffolding or instructional support from their teacher or peers.

However, the superior level of social reasoning revealed in the CSR group implies that early adolescents need effective instructional support to be able to overcome certain limitations in their social reasoning. The argumentation literature has documented several limitations of reasoning in students and adult learners. For example, there is evidence that students have to make substantial efforts to overcome their biased prior beliefs when they evaluate evidence, a phenomenon called myside or confirmation bias (e.g., Villarroel, Felton, & Garcia-Mila, 2016). Other studies have found that students often fail to consider counterarguments in their own writing (e.g., Nussbaum et al., 2005). Even if counterarguments are considered, students are likely to encounter difficulties in weighing and integrating multiple perspectives (Nussbaum & Schraw, 2007).

Some of these limitations were present in our students' social reasoning essays. In the pre-test, when asked to reason about social exclusion/injustice issues, the students from both conditions tended to draw more information from their personal and moral knowledge than from their societal knowledge or based on pragmatic considerations. Specifically, a significant number of reasons were generated by students in the pre-test to support one side of the social exclusion issue; these reasons were associated with the story characters' personal concerns or benefits (e.g., gaining/losing friends, social status, physical appearance) and moral principles (e.g., it is not right to do harm to others physically or verbally). Even though these reasons remained salient in the post-test essays, the CSR group students were more likely to extend their reasoning to societal rules or pragmatic consequences of a social action than those in the RA group. The more coherent, complex, and coordinated social reasoning essays generated by the CSR group further suggest that these students became more competent in making social-moral decisions based on cross-domain coordination (Nucci, 2009).

The CSR students' superior ability to consider the complexity of a social-moral issue was also supported by the greater number of responses or solutions that they generated in the essays. The big questions in the social reasoning essay task asked students to make a dichotomous social-moral decision (i.e., whether the protagonist should perform an action or not) for the purpose of minimizing any confusion that the fifth-grade students might have as they wrote. When encountering such a dichotomous question, students could easily narrow their final response to the big question

to one of the dichotomous options. Our findings showed that the CSR group was more likely to think beyond these dichotomous solutions than the RA group.

Another major finding was the quality of reasoning revealed in the essays by the CSR group compared to the RA group. Even though the number of interpretations was greater in the RA group after controlling for pre-test scores, the two groups differed in the number of shallow analogies. An example is that some students made a comparison between physical aggression mentioned in the story to their own personal experience with physical aggression, but soon digressed to speaking of their own stories without referencing back to their main arguments. This finding suggests that students who experienced CSR discussions were more able to stay on topic with their central arguments.

In the CSR discussions, students were asked to take both critical-analytic and expressive stances while reasoning about social exclusion and injustice. Students in the CSR group were encouraged to rationally evaluate different reasons and evidence drawn from the story and other sources of information. They were also encouraged to share relevant emotional experiences as a way to construe story characters' affective states and social motives. The rationale of this design was that students would learn to engage in both 'hot' cognition (i.e., reasoning with strong feelings or high arousal) and 'cold' cognition (i.e., reasoning with low emotion or arousal). Interestingly, the finding of social perspective taking did not fully support our hypothesis that CSR students would consider more cognitive and affective perspectives of the story characters than the RA students. Students from the CSR condition became more attuned to cognitive perspective taking, whereas students from the RA condition generated more affective perspective taking. One of Haidt's (2007) principles of morality called Intuitive Primacy stressed that the human mind is wired to process affective information more automatically and immediately than cognitive information. Compared to the RA condition, the CSR discussions seem to have lowered the degree to which students generated such intuitive responses in their essays. Our finding thus calls for a deeper understanding of students' use of affective sources of information in their arguments.

One study limitation was the unequivalent pre-test social reasoning scores of the two groups of students. This limited our ability to infer students' growth of social reasoning. Students' pre-test scores might have been affected by the study's consent process. Classrooms were assigned to conditions prior to the student consent process. Following the IRB protocol, students were given different consent forms based on the conditions to which they were assigned. This consent process might have heightened CSR students' motivation to perform well in the pre-test compared to the RA group. However, by controlling for pre-test social reasoning performance, our findings could only possibly underestimate the effects of the CSR discussions.

The study did not include a business-as-usual control comparison group, which limited our ability to account for story effects as one potential confounding factor of the intervention. In both the CSR and RA conditions, the students were exposed to the same story materials once a week for five consecutive weeks. It is likely that changes in their social reasoning from the pre- to the post-test were attributed to the stories they read. Since the five stories chosen for the current study shared the common theme of social exclusion and injustice, students might have been able to inductively generate social knowledge through inter-textual comparisons. Without a business-as-usual control group, less is known about whether the social reasoning ability observed in the

CSR group was mainly attributed to the effect of dialogic inquiry or a combination of dialogic inquiry and story effects. However, some studies have reported the positive influence of dialogic instruction on students' reasoning compared to traditional lecturing (Chinn et al., 2001). Comparing the CSR approach with the RA approach in which students were systematically introduced to the same story materials through RA and individual writing remains a superior study design for inferring the causal influence of collaborative dialogic inquiry on students' social reasoning.

Despite the positive influence of the CSR discussions on individuals' social reasoning, this intervention incorporated other instructional elements to cultivate an intellectually stimulating environment for dialogic inquiry, and to ensure the ecological validity of the intervention. These instructional strategies, such as goal setting and teacher's prompts, have been shown to foster student engagement in dialogic inquiry (e.g., Nussbaum et al., 2005; Webb, Franke, Johnson, Ing, & Zimmerman, 2019). However, it was impossible for us to examine the potential influences of these instructional elements on the student outcome in one single study. Future research is needed to further examine the unique contributions of these instructional elements.

Educational implications and scientific contributions

In response to a recent educational initiative to incorporate social-emotional learning in formal education, the study proposed an evidence-based instructional approach that can effectively enhance students' social reasoning. Collaborative Social Reasoning discussions are effective in helping students overcome their cognitive limitations through a deliberative, non-egocentric process of argumentation, as well as an equitable and supportive process of intellectual collaboration. The current findings contribute to the fields of moral development and argumentation with a deeper understanding of how dialogic interaction can be a vehicle to expand students' social knowledge, to deepen social information processes, and to seek a balance between intuitive and rational models of reasoning.

Acknowledgements This research was supported by Chiang Chin-Kuo Foundation and the Institute for Research Excellence in Learning Sciences of National Taiwan Normal University (NTNU) from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan. The opinions expressed are those of the authors and do not necessarily represent views of the foundation. We would like to thank the research team and the many administrators, teachers, and children without whom this study would not have been possible.

Appendix: Chinese version of the three sample essays

Essay 1

如果我是大衛的話我會先確認是誰做的, 如果不是布蘭登我就會通知校長有人寫了這份貼文, 如果是布蘭登的話我就會先勸他刪除這份貼文, 如果他不肯我就會告訴他如果被人發現會有什麼後果, 他還是認爲很好玩不想刪除的話,我會立即告發他, 雖然一定會影響到我的人際關係, 但是我們應該做對的事, 這和

我沒有關係，但是這是正義的舉動我們應該做正義的事，可是我也有可能直接不理這張照片當作沒有發生過，因爲我不想得罪我珍貴的好友大家都說薩曼紗是老師的馬屁精，他就是一個非常討厭的人，我不會認爲這是對的，但我會保密因爲我跟他不熟也不是朋友。

Essay 2

我覺得大衛應該要告訴校長這則訊息,這樣可以防止讓他的那些同學繼續傳那種貼文，這實在很沒品德,而且傳這張照片的「那個人」，如果是別人傳他的照片，我想看到照片時一定也很難過,妳讓別人難過很簡單，但是讓他人開心很難，即使妳很瞭解他，妳會覺得他的外表看起來很開心，內心呢?妳有沒有想過，搞不好他的內心是刺痛、很難的，妳想像不出來的痛,所以我們要多多爲人著想，不要去看那些不好的東西，爲人處世，不可以袖手旁觀，就能共享友情天地!

Essay 3

我認爲大衛應該跟老師和師長門說有關於這則貼文的事情，因爲布蘭登和其他人把薩曼紗的照片塗滿螢光色，還在社群上面嘲笑他，還用下流的語言文字去污辱他，而且薩曼紗並沒有對他們做什麼讓他們不舒服的事情，這樣也代表布蘭登和其他人是惡意要限制他人的自由，就像逼迫他人一定要照他的方法做一樣，根據憲法的自由權這部份,他們都已經犯法了，薩曼紗有參與和個人的自由和隱私權,他可以想要做什麼就做什麼，重要的是他並沒有妨礙到他人，布蘭登和其他人應該事先想到「這樣做,別人會不會不舒服」，　而他們也該要想到將心比心，要好好對待他人，他人也就會好好對待妳，可是他們並沒有想到那麼多，他們只覺得好玩就好,好笑就好，不用顧慮到其他人，所以他們不但做出了語言罷凌、妨礙自由權、侵犯隱私權等罪。

我認爲大衛該舉發他們，而大衛不該因爲怕會沒有朋友，就把最重要的事情放在一旁都不說，等到真的有一天有人去舉發了，這樣大衛也就變成共犯，這樣雙方都沒有好處，也算是「兩敗俱傷」了，所以大衛應該要儘早去舉發這則貼文。

References

Alexander, R. J. (2006). *Towards dialogic teaching: Rethinking classroom talk* (3rd ed.). York: Dialogos.

Bengtsson, H., & Arvidsson, Å. (2011). The impact of developing social perspective-taking skills on emotionality in middle and late childhood. *Social Development, 20,* 353–375. https://doi.org/10.1111/j.1467-9507.2010.00587.x.

Berkowitz, M. W., Gibbs, J. C., & Broughton, J. M. (1980). The relation of moral judgment stage disparity to developmental effects of peer dialogues. *Merrill-Palmer Quarterly of Behavior and Development, 26,* 341–357. https://doi.org/10.2307/23084042.

Chen-Gaddini, M. (2012). Chinese mothers and adolescents' views of authority and autonomy: A study of parent–adolescent conflict in urban and rural China. *Child Development, 83,* 1846–1852. https://doi.org/10.1111/j.1467-8624.2012.01823.x.

Cheung, C. K., & Lee, T. Y. (2010). Contributions of moral education lectures and moral discussion in Hong Kong secondary schools. *Social Psychology of Education, 13,* 575–591. https://doi.org/10.1007/s11218-010-9127-x.

Chinn, C. A., Anderson, R. C., & Waggoner, M. A. (2001). Patterns of discourse in two kinds of literature discussion. *Reading Research Quarterly, 36,* 378–411. https://doi.org/10.1598/RRQ.36.4.3.

Cohen, E. G., & Lotan, R. A. (1995). Producing equal-status interaction in the heterogeneous classroom. *American Educational Research Journal, 32,* 99–120. https://doi.org/10.2307/1163215.

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin, 115,* 74–101. https://doi.org/10.1037/0033-2909.115.1.74.

Crick, N. R., & Dodge, K. A. (1999). 'Superiority is in the eye of the beholder: A comment on Sutton, Smith, and Swettenham. *Social Development, 8,* 128–131. https://doi.org/10.1111/1467-9507.00084.

Dahl, A., & Kim, L. (2014). Why is it bad to make a mess? Preschoolers' conceptions of pragmatic norms. *Cognitive Development, 32,* 12–22. https://doi.org/10.1016/j.cogdev.2014.05.004.

Damon, W. (1980). Structural-developmental theory and the study of moral development. In M. Windmiller, N. Lambert, & E. Turiel (Eds.), *Moral development and socialization* (pp. 35–68). Boston, MA: Allyn and Bacon.

Damon, W., & Killen, M. (1982). Peer interaction and the processes of change in children's moral reasoning. *Merrill-Palmer Quarterly, 28,* 347–367.

Denison, C. (2014). Ethical dilemmas for classroom discussion: The daily dilemma #29. *Good Character. com*. Retrieved from http://www.goodcharacter.com/dilemma/dilemma29.html.

Diazgranados, S., Selman, R. L., & Dionne, M. (2016). Acts of social perspective taking: A functional construct and the validation of a performance measure for early adolescents. *Social Development, 25,* 572–601. https://doi.org/10.1111/sode.12157.

Fitzhugh, L. (2009). *Harriet the spy*. New York, NY: Harper & Row.

Fraser, M. W., Galinsky, M. J., Smokowski, P. R., Day, S. H., Terzian, M. A., Rose, R. A., et al. (2005). Social information-processing skills training to promote social competence and prevent aggressive behavior in third grade. *Journal of Consulting and Clinical Psychology, 73,* 1045–1055. https://doi.org/10.1037/0022-006X.73.6.1045.

Gibbs, J. C. (2003). *Moral development and reality: Beyond the theories of Kohlberg and Hoffman, and Haidt*. Thousand Oaks, CA: Sage.

Golanics, J. D., & Nussbaum, E. M. (2008). Enhancing collaborative online argumentation through question elaboration and goal instructions. *Journal of Computer Assisted learning, 24,* 167–180. https://doi.org/10.1111/j.1365-2729.2007.00251.x.

Haidt, J. (2007). The new synthesis in moral psychology. *Science, 316*(5827), 998–1002. https://doi.org/10.1126/science.1137651.

Howe, A. (1990). A climate for small group talk. In M. Brubacher, R. Payne, & K. Rickens (Eds.), *Perspectives on small group learning* (pp. 101–118). Oakville, CA: Rubicon.

Johnson, D. W. (1975). Cooperativeness and social perspective taking. *Journal of Personality and Social Psychology, 31,* 241–244. https://doi.org/10.1037/h0076285.

Kinney, J. (2013). *Diary of a wimpy kid*. New York, NY: Amulet Books.

Kohlberg, L. (1973). The claim to moral adequacy of a highest stage of moral judgment. *The Journal of Philosophy, 70,* 630–646. https://doi.org/10.2307/2025030.

Kohlberg, L., & Hersh, R. (1977). Moral development: A review of the theory. *Theory into Practice, 16,* 53–59. https://doi.org/10.1080/00405847709542675.

Lansford, J. E., Malone, P. S., Dodge, K. A., Pettit, G. S., & Bates, J. E. (2010). Developmental cascades of peer rejection, social information processing biases, and aggression during middle childhood. *Development and Psychopathology, 22,* 593–602. https://doi.org/10.1017/S0954579410000301.

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

Lin, T.-J., Anderson, R. C., Jadallah, M., Nguyen-Jahiel, K., Kim, I.-H., Kuo, L.-J., et al. (2015). Social influences on children's development of relational thinking during small-group discussions. *Contemporary Educational Psychology, 41,* 83–97. https://doi.org/10.1016/j.cedpsych.2014.12.004.

Loban, W. (1976). *Language development: Kindergarten through grade twelve (No. 18).* Urbana, IL: National Council of Teachers of English.

Lourenço, O. (2014). Domain theory: A critical review. *New Ideas in Psychology, 32,* 1–17.

Ma, S., Anderson, R. C., Lin, T.-J., Zhang, J., Morris, J. A., Nguyen-Jahiel, K., et al. (2017). Instructional influences on English language learners' storytelling. *Learning and Instruction, 49,* 64–80. https://doi.org/10.1016/j.learninstruc.2016.12.004.

McVee, M. B. (2014). "Some are way left, like this guy, Gloria Ladson-Billings": Resistance, conflict, and perspective taking in teachers' discussions of multicultural education. *Peace and Conflict: The Journal of Peace Psychology, 20,* 536–551. https://doi.org/10.1037/pac0000067.

Mercer, N., & Dawes, L. (2008). The value of exploratory talk. In N. Mercer & S. Hodgkinson (Eds.), *Exploring talk in school*. London: Sage.

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

Mischo, C. (2005). Promoting perspective coordination by dilemma discussion. The effectiveness of classroom group discussion on interpersonal negotiation strategies of 12-year-old students. *Social Psychology of Education, 8,* 41–63.

Morris, J. A., Miller, B. W., Anderson, R. C., Nguyen-Jahiel, K., Lin, T.-J., Scott, T., et al. (2018). Instructional discourse and argumentative writing. *International Journal of Educational Research., 90,* 234–247. https://doi.org/10.1016/j.ijer.2018.03.001.

Mulvey, K. L. (2016). Children's reasoning about social exclusion: Balancing many factors. *Child Development Perspectives, 10,* 22–27. https://doi.org/10.1111/cdep.12157.

Murphy, P. K., Wilkinson, I. A., Soter, A. O., Hennessey, M. N., & Alexander, J. F. (2009). Examining the effects of classroom discussion on students' comprehension of text: A meta-analysis. *Journal of Educational Psychology, 101,* 740–764. https://doi.org/10.1037/a0015576.

Myyrya, L., Juujärvi, S., & Pesso, K. (2010). Empathy, perspective taking and personal values as predictors of moral schemas. *Journal of Moral Education, 39,* 213–233. https://doi.org/10.1080/03057241003754955.

Nucci, L. (2009). *Nice is not enough: Facilitating moral development*. Saddle River, NJ: Merrill.

Nucci, L., Creane, M. W., & Powers, D. W. (2015). Integrating moral and social development within middle school social studies: A social cognitive domain approach. *Journal of Moral Education, 44,* 479–496. https://doi.org/10.1080/03057240.2015.1087391.

Nussbaum, E. M., Kardash, C. M., & Graham, S. E. (2005). The effects of goal instructions and text on the generation of counterarguments during writing. *Journal of Educational Psychology, 97,* 157–169. https://doi.org/10.1037/0022-0663.97.2.157.

Nussbaum, E. M., & Schraw, G. (2007). Promoting argument-counterargument integration in students' writing. *The Journal of Experimental Education, 76,* 59–92. https://doi.org/10.3200/JEXE.76.1.59-92.

Oberle, E., Schonert-Reichl, K. A., & Thomson, K. C. (2010). Understanding the link between social and emotional well-being and peer relations in early adolescence: Gender-specific predictors of peer acceptance. *Journal of Youth and Adolescence, 39,* 1330–1342. https://doi.org/10.1007/s10964-009-9486-9.

Palacio, R. J. (2012). *Wonder*. New York, NY: Alfred A. Knopf.

Park, Y., & Killen, M. (2010). When is peer rejection justifiable? Children's understanding across two cultures. *Cognitive Development, 25,* 290–301. https://doi.org/10.1016/j.cogdev.2009.10.004.

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

QSR International Pty Ltd. (2015). NVivo qualitative data analysis Software (Version 11).

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

Reznitskaya, A., Kuo, L., Clark, A., Miller, B., Jadallah, M., Anderson, R. C., et al. (2009). Collaborative reasoning: A dialogic approach to group discussions. *Cambridge journal of education, 39,* 29–48. https://doi.org/10.1080/03057640802701952.

Ripken, C., Jr., & Cowherd, K. (2012). *Supper-sized slugger*. New York, NY: Disney Book Group.

Sachar, L. (1987). *There is a boy in the girls' bathroom*. New York, NY: Yearling Books.

Smetana, J. G. (2006). Social domain theory: Consistencies and variations in children's moral and social judgments. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 119–154). Mahwah, NJ: Erlbaum.

Smith, R. L., & Rose, A. J. (2011). The "cost of caring" in youths' friendships: Considering associations among social perspective taking, co-rumination, and empathetic distress. *Developmental Psychology, 47,* 1792. https://doi.org/10.1037/a0025309.

Sun, J., Anderson, R. C., Lin, T.-J., & Morris, J. A. (2015). Social and cognitive development during Collaborative Reasoning. In L. B. Resnick, C. Asterhan, & S. N. Clarke (Eds.), *Socializing intelligence through academic talk and dialogue* (pp. 63–76). Washington, DC: American Educational Research Association.

Sun, J., Anderson, R. C., Perry, M., & Lin, T.-J. (2017). Emergent leadership in children's cooperative problem solving groups. *Cognition and Instruction, 35,* 212–235. https://doi.org/10.1080/07370008.2017.1313615.

Sutton, J., Smith, P. K., & Swettenham, J. (1999). Bullying and 'theory of mind': A critique of the 'social skills deficit's view of anti-social behaviour. *Social Development, 8,* 117–127. https://doi.org/10.1111/1467-9507.00083.

Tisak, M. (1993). Preschool children's judgments of moral and personal events involving physical harm and property damage. *Merrill-Palmer Quarterly, 39,* 375–390.

Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge, UK: Cambridge University Press.

Turiel, E. (2006). Thought, emotions, and social interactional processes in moral development. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 7–36). Mahwah, NJ: Erlbaum.

Villarroel, C., Felton, M., & Garcia-Mila, M. (2016). Arguing against confirmation bias: The effect of argumentative discourse goals on the use of disconfirming evidence in written argument. *International Journal of Educational Research, 79,* 167–179. https://doi.org/10.1016/j.ijer.2016.06.009.

Walker, C. L., Shore, B. M., & Tabatabai, D. (2013). Eye of the beholder: Investigating the interplay between inquiry role diversification and social perspective taking. *International Journal of Educational Psychology, 2,* 144–192. https://doi.org/10.4471/ijep.2013.23.

Walton, D. (1998). *The new dialectic: Conversational contexts of argument*. Toronto: University of Toronto Press.

Webb, N. M., Franke, M. L., Johnson, N. C., Ing, M., & Zimmerman, J. (2019). Promoting productive student participation across multiple classroom participation settings. In R. Gillies (Ed.), *Promoting academic talk in schools: Global practices and perspectives*. Abingdon, OX: Taylor and Francis.

Wells, G. (1999). *Dialogic inquiry: Toward a sociocultural practice and theory of education*. Cambridge: Cambridge University Press.

Wolff, V. E. (2000). *Bat 6*. New York, NY: Scholastic.

Zhang, X., Anderson, R. C., Dong, T., Nguyen-Jahiel, K., Li, Y., Lin, T.-J., et al. (2013). Children's moral reasoning: Influence of culture and collaborative discussion. *Journal of Cognition and Culture, 13,* 503–522. https://doi.org/10.1163/15685373-12342106.

Zhang, X., Anderson, R. C., Morris, J., Miller, B., Nguyen-Jahiel, K. T., Lin, T. J., et al. (2016). Improving children's competence as decision makers: Contrasting effects of collaborative interaction and direct instruction. *American Educational Research Journal, 53,* 194–223. https://doi.org/10.3102/0002831215618663.