Educ Psychol Rev (2015) 27:79–102
DOI 10.1007/s10648-014-9260-8

REVIEW ARTICLE

Effects of Small-Group Learning on Transfer: a Meta-Analysis

Hui-Hua Pai · David A. Sears · Yukiko Maeda

Published online: 23 February 2014

Abstract This study investigated the potential benefit of small-group learning on transfer performance using the method of meta-analysis. Results showed positive support for the hypothesis that small-group learning can increase students' transfer performance (average effect size of 0.30). Unlike reviews of effects of cooperation on learning, this review of effects on transfer found no greater benefit of structured small-groups compared to unstructured. This finding, in conjunction with the significant variability found across effect sizes, suggests that further investigation into features of effective unstructured small-group tasks might yield heuristics that teachers could eventually use to make decisions about when collaboration would be most useful. Although some of the reviewed studies were published decades ago, the vast majority were published within the last few years, suggesting a growing area of research interest.

Keywords Small-group learning · Cooperative learning · Collaborative learning · Transfer · Meta-analysis

Having students work in small groups to achieve shared learning goals has been recommended by researchers for transforming students' learning experiences and outcomes (Blumenfeld

Dr. Pai and Dr. Sears contributed equally to this manuscript

H.-H. Pai
Beering Hall of Liberal Arts and Education, Purdue University, Room 5123, 100 N. University Street, West Lafayette, IN 47907-2067, USA
e-mail: hpai@purdue.edu

D. A. Sears
Beering Hall of Liberal Arts and Education, Purdue University, Room 5130, 100 N. University Street, West Lafayette, IN 47907-2067, USA

Y. Maeda
Beering Hall of Liberal Arts and Education, Purdue University, Room 5152, 100 N. University Street, West Lafayette, IN 47907-2067, USA
e-mail: ymaeda@purdue.edu

H.-H. Pai · D. A. Sears (✉)
Purdue University, BRNG 5130, 100 N. University Ave, West Lafayette, IN 47907, USA
e-mail: dsears@purdue.edu

et al. 1996). Small-group learning has shown better effects on students' academic-related outcomes than competitive and individualistic learning methods (Johnson et al. 1981). Structures, such as scripts, roles, and group rewards, have been identified as critical for fostering greater learning in groups than in individual contexts (e.g., Johnson and Johnson 2009; Johnson et al. 2000; O'Donnell 2006; Slavin 1996). While small-group learning has shown tremendous success at promoting students' acquisition of knowledge, its effects on their transfer of knowledge is less clear. Transfer is the ability to apply or adapt prior knowledge to a novel situation (Bransford et al. 2000). A growing number of studies have recently identified benefits to transfer from collaboration (e.g., Kirschner et al. 2009; Sears and Pai 2012). The current study provides a formal examination of the burgeoning empirical relationship between small-group learning and transfer in the form of a meta-analysis.

Small-Group Learning

Small-group learning is used here to encompass contexts in which students work together in small groups to achieve shared learning goals (Dillenbourg 1999; Johnson and Johnson 1999). This definition is similar to O'Donnell's (2006) use of "peer learning" (p. 781) in that it includes both cooperative learning, collaborative learning, and reciprocal teaching/learning formats such as jigsaw (Aronson 2002). We occasionally use the terms cooperative, collaborative, and small-group learning interchangeably. While cooperative and collaborative learning have different origins and the former is often more structured, both involve students working together to achieve shared learning goals, both have a history of successfully increasing student learning relative to individual study, and the terms are similar enough that they are frequently used interchangeably (Bruffee 1995; Johnson and Johnson 1996; Manion and Alexander 1997; Matthews et al. 1995; O'Donnell 2006; Panitz 1999).

Small groups usually comprise two to four or five students (Johnson and Johnson 2009; O'Donnell 2006; Slavin 1977, 1995). Some debate exists about whether groups of only two people should count as a group, especially because certain social phenomena are not possible without at least three people, such as ostracization (cf. Moreland 2010 vs. Williams 2010). Rau and Heyl (1990) note that as group size increases, achieving productive group coordination and meaningful participation by all members becomes more complex. Nevertheless, many of the extensively studied small-group learning phenomena involve dyadic exchanges, such as one student summarizing or explaining while the other listens or checks for errors (e.g., O'Donnell et al. 1987; Webb 1982). While variations in small-group size may play important roles in some collaborative tasks, the focus in the current study is on small groups in general.

Research has shown positive effects of small-group learning on students' academic success (e.g., Johnson et al. 1998; Johnson and Johnson 1999; O'Donnell et al. 1987; O'Donnell 2006; Slavin 1996). The factors contributing to its effectiveness in promoting student achievement have gained attention in recent decades. External scaffolds or structures, such as scripts, roles, and reward systems, have been found to be particularly important for promoting significant gains in group learning compared to individual study. For example, Slavin (1983, 1996) found group goals and individual accountability to be defining features of effective cooperative learning. These motivational structures create an environment in which individuals can only achieve their personal goals when the group succeeds—a condition known as positive interdependence (Johnson and Johnson 1999, 2009). Rewarding a group for doing well can induce goal-oriented behaviors, and individual accountability can ensure every member is contributing and engaged. Methods providing group rewards without individual accountability to those goals might succumb to social loafing where only one or two members do most of the work (Latané et al. 1979; Michaelsen et al. 1997). In contrast, if individual grades or feedback are given but no group goals applied, students might not engage in behaviors to help their teammates succeed. In Slavin's (1995, 1996) review of this topic, studies incorporating both group goals and individual accountability yielded a higher median effect size (+0.32) than studies lacking either component (+0.07). Johnson et al.'s (2000) meta-analysis involving 164 studies of kindergarteners through adults from 1970 to 1999 revealed similar benefits of cooperative learning methods that included positive interdependence. Effect sizes of these different methods of cooperative learning compared to individualistic learning ranged from 0.13 to 1.04.

Carefully structured interaction fostered by training students in particular collaborative skills is another approach that has shown benefits to group learning (Gillies 2004). For example, Ashman and Gillies (1997) compared school-age students who were trained to practice interpersonal skills (e.g., clear expression of ideas, active listening skills and giving constructive criticism) and cooperative skills (e.g., sharing tasks evenly, taking turns, and resolving different opinions) to their untrained peers. The trained students obtained higher posttest results in the social studies activities than the untrained groups. Additionally, the trained groups were more responsive to their peers' needs, and showed more positive attitudes toward the activities. These differences were maintained over time.

Effective small-group learning interaction can also be structured by scripts that are used to specify the roles of group members or the sequences of activities they engage in. Scripts are designed to increase specific cognitive behaviors associated with learning, such as summarizing, providing explanations, or asking questions (e.g., Coleman 1998; O'Donnell 1999; Palincsar and Brown 1984). One example of scripted cooperation is provided by O'Donnell et al. (1987). After reading a section of text, one partner of the scripted dyad summarized the material and the other detected errors or omissions in the summary, and then they switched roles. Students who use scripted cooperation often perform better and show positive affect toward their partners and the tasks (O'Donnell 1999).

Task specialization, where students within a group each become experts on related subtopics of a larger project, is another method by which small-group learning is structured (Slavin 1996). This method can create positive interdependence among group members when each member must learn from their peers in order for the group to succeed on the task. Aronson's Jigsaw method is one of the cooperative models implementing task specialization (Aronson 2002; Aronson et al. 1978). Each group member studies a subtopic of the material, meeting in "expert groups" to share information with peers from the other jigsaw groups specializing in the same subtopic, and then returning to their groups to teach their peers about their subtopic. Each student is like a piece in a jigsaw puzzle. Each part is essential for full understanding of the final product (Aronson 2002). In Sharan and Sharan's (1992) Group Investigation method, students take an active role in planning how to study a topic and how to form a cooperative group. They divide the work among themselves and each group member is responsible for investigating his or her part of the topic. Ultimately, the group synthesizes their findings and presents their work to the whole class. These methods emphasize teambuilding processes meant to motivate students to help each other succeed. When group rewards are applied, to further support the alignment of group and individual goals, the cooperative learning methods that emphasize task specialization can be very effective for student achievement (Slavin 1996).

All the aforementioned factors provide structures to scaffold small-group learning to work to its full potential. However, structure is not the only factor that affects the effectiveness of collaboration. The nature of the task can also play a role. A number of studies have shown no benefit of unstructured collaboration when a task is simple. For example, Andersson and

Rönnberg (1995) examined dyads' recall in episodic memory tasks. College students who worked in dyadic groups suffered in the single-word free recall task. Social interaction in the study seemed to inhibit students from retrieving the items recalled previously. Lamm and Trommsdorff's (1973) review of group performance on brainstorming activity also demonstrated a disadvantage of using groups for idea-generation tasks. Working alone and then combining the results with other peers produced more numbers of ideas than the same number of people cooperating in groups. Lamm and Trommsdorff provided several possible interpretations of this finding. One was production blocking—only one group member can speak at a time during group brainstorming. Also, participants may hold back ideas due to the fear of being judged negatively by other people. Social inhibition makes participants withhold ideas out of fear of criticism, and "one-track thinking" can further hinder group performance (Lamm and Trommsdorff 1973, p. 382).

In contrast, some tasks have shown benefit from collaboration even when the external structuring is limited. Barron (2000) assessed sixth graders' problem solving ability by requiring them to generate plans and quantitative solutions to a series of problems in a video. Students who worked in triads outperformed students in the individual condition on mastery and transfer problems. While this study did involve supports for student learning, such as asking students to identify key data and procedures before asking them to construct a complete solution, it did not specify their interaction processes with added external supports such as scripts, roles, and rewards. In other words, it was a task where positive interdependence and productive interaction behaviors were not promoted by external structures.

In the study of Laughlin et al. (2008), a challenging letters-to-numbers decoding problem was used to assess college students' problem solving and reasoning skills. The result indicated that groups performed better than individuals, and this superior performance subsequently transferred from group to individual performance. Schwartz (1995) compared the cognitive products generated by groups to the ones produced by individuals when solving novel problems. He found that dyads were significantly more likely to induce numerical rules and abstract visualizations than individuals. Further, he concluded that when working in groups, multiple perspectives on the problem need to be negotiated to a common representation. Therefore, the representation tends to be abstract to be able to bridge various views. Collaboration provides an environment to generate more abstract representations which is not normally available when working alone. The positive effects of collaboration on complex tasks, even without external structures, suggest that one of the unique benefits of collaboration to learning may be its ability to promote deep understanding or transfer.

Small-Group Learning and Transfer

Transfer is the ability to apply or adapt knowledge to novel contexts (e.g., Bransford et al. 2000; Perkins and Salomon 1992; Schwartz et al. 2008). It is one of the imperative goals of education to facilitate performance and learning in contexts inside and outside of the classroom (Barnett and Ceci 2002; Bransford and Schwartz 1999). However, transfer is not easy to achieve. Novices tend to attend to surface features and struggle to extract deep structure that enables knowledge to transfer positively (Chi et al. 1981; Novick 1988). Even when people have relevant prior knowledge, they often need to be cued to use it in a transfer context (Gick and Holyoak 1980, 1983). Despite over a century of research, transfer remains difficult to predict and promote (Lobato 2006).

Small-group learning has been shown to promote knowledge acquisition and comprehension (e.g., Lou et al. 2001), abstraction (e.g., Schwartz 1995), complex problem-solving (e.g., Sears and Reagin 2013), higher levels of metacognitive thinking (e.g., Manion and Alexander 1997), as well as motivational and social benefits (e.g., Sears and Pai 2012; Slavin 1996; Slavin and Cooper 1999). Learning with other people enables the interaction and negotiation process thought to promote deep understanding that working in isolation often fails to facilitate. Group members need to resolve different perspectives of the problem and bring multiple views together. Developing a shared view may require finding less idiosyncratic and more abstract representations of a problem (Schwartz 1995). It also promotes giving and receiving explanations—processes known to promote deeper understanding of materials (Chi et al. 1994; Roscoe and Chi 2008; Webb 1982). Group members may encourage and motivate each other to increase effort or persistence (Sears and Pai 2012; Slavin 1983). These components of collaboration create a learning environment that invites learners to engage deeper cognitive abilities to comprehend materials, construct abstract representations, and promote metacognitive thinking.

Greater initial understanding, abstraction, and metacognition are associated with greater transfer (Bransford et al. 2000; Brown 1989; Chi and VanLehn 2012; Gick and Holyoak 1983). Individuals need to have learned the original subject sufficiently well to provide a basis for further application. Small-group learning creates a unique environment to enrich initial understanding (e.g., Cohen 1994; Johnson et al. 2000; O'Donnell 2006; Slavin 1996). While working collaboratively, individuals have to generate and explain their thoughts to each other. Vocalizing one's thoughts can help to produce an organized cognitive structure of the material. Especially when vocalizing to a peer, students may view themselves as teachers and may exert greater effort towards clarity of concepts, so others may understand (Webb 1982). They may also have their own thoughts reflected back to them from their peer, a form of recursive feedback that can support learning (Okita and Schwartz 2013). Through interaction, group members can reconstruct their ideas and form new understanding they may not have discovered by themselves (Webb 1982).

Transfer is enhanced when learners have elicited critical attributes of a situation and developed the ability to flexibly deploy key principles (Biederman and Shiffrar 1987; Bransford and Schwartz 1999; Schwartz et al. 2005). Abstraction and differentiation are ways to discern the key features and deep structures of a domain and thereby avoid knowledge being too context-bound (Gick and Holyoak 1980, 1983; Marton 2006; Schwartz and Bransford 1998; Singley and Anderson 1989). Both of these processes tend to involve actively comparing and contrasting different examples in order to distinguish surface features from deep features. Transfer requires understanding beyond the superficial elements of the material. Working in groups encourages students to exchange ideas and reconcile competing opinions. Students can compare multiple perspectives and examples the group members bring to the discussion. Comparing the similarities and differences between various examples pushes students to notice and differentiate contrasting features (Bransford and Schwartz 1999; Gick and Holyoak 1983; Marton 2006; Schwartz and Bransford 1998). Group members may abstract a common representation so they can reconcile conflicting perspectives or negotiate meaning from common ground (Schwartz 1995). These collaborative processes should engage students to filter the surface features and extract the deep structures from the materials. Recent work has suggested that students in small-groups may be better at discerning and transferring key concepts than individuals (Sears et al. 2011).

Tasks which require more effortful and cognition-engaged processing—as opposed to tasks that require less active transformation of inputs, such as copying or rote memorization—are associated with greater transfer (e.g., Brown 1989; Needham and Begg 1991; Phelps and Damon 1989). Transfer can occur automatically, but it often does not and requires, instead, metacognitive awareness, motivation, and intentional search for prior knowledge that may solve a current problem (Perkins and Salomon 2012; Salomon and Perkins 1989). Compared to tasks only involving memorization and following simple procedures, students tend to use more strategies to regulate their learning when working on complex transfer tasks (Perkins and Salomon 1992). By practicing self-regulated learning, students gain more practice identifying what strategies they can use when encountering novel situations. Peer collaboration has shown benefits on strategy use and metacognitive causal attribution (Manion and Alexander 1997). Children with lower levels of metacognition were able to increase their use of sorting strategies and induce higher levels of metacognitive thinking when learning in groups that included students with more metacognitive sophistication (Manion and Alexander 1997). The researchers found that group members were observing and mimicking others' behaviors.

As suggested by previous research (e.g., Cohen 1994; Cohen et al. 1999) and by the cognitive processes supported by group interaction, small-group learning may be particularly supportive of transfer. Kirschner et al. (2009) provide an illustrative example of this point. Seventy high school students were randomly assigned to individual or group conditions to learn about heredity in the domain of biology. They received three learning tasks to solve problems about the proportion of possible genotypes of offspring. The students in the individual condition received all information while each member in the triadic groups only received one third of the information to solve the problems. Subsequently, three retention tests that were nearly identical to the training tasks, and three transfer tests that were structurally different from the learning tasks were given. The result showed that students who learned in a group outperformed the individual learners on the transfer tasks but not on the retention tasks. Kirschner et al. noted that working in a group enables learners to take advantage of sharing the cognitive load and having more cognitive capacity left to comprehend the information at a deeper level instead of merely remembering it. This differentiation of transfer from retention highlights why it is important to determine if the effects of small-group learning might be different for transfer than they are for learning in general, both in terms of the magnitude of the typical effect and in terms of the scaffolds that may (or may not) be required for those effects.

Present Study

The main purpose of this study was to synthesize the results of existing empirical studies of small-group learning effects on transfer by using the meta-analytic method. The effect size of interest in each included study was defined as the standard mean difference between the transfer performance in small group learning and that in individual learning. Positive effect sizes indicate superiority of transfer performance for groups relative to individuals. The key outcome of interest was the average effect of small-group learning on students' transfer performance, including the direction and magnitude of the effect across studies. Variation of the effect sizes across studies was investigated, and moderator analysis with a key theory-based variable was conducted to attempt to explain the variation. Using external structures, such as group goals and individual accountability or scripts and roles, has been effective for promoting learning in small groups (Johnson and Johnson 1999; O'Donnell 2006; Slavin 1996). Therefore, structured and unstructured small-group learning served as the theory-based moderator. Finer-grained distinctions between different types of structures, such as Slavin's (1996) comparison of different incentive structures, are important but were not feasible in the current study due to limited numbers of studies explicitly examining transfer in each structure sub-category.

Based on previous research, small-group learning was hypothesized to show superior transfer performance compared to individual learning. Studies of small-group learning come from different theoretical traditions and often utilize unique designs, including different collaboration methods, learning tasks, and assessment tests. Effect sizes across studies were expected to show significant variability. In reviews of the effects of small groups on learning, unstructured group work has often proven ineffective at promoting learning whereas structured methods have yielded typical effect sizes of 0.32 (Slavin 1996). In reviews of other structured methods, similarly positive results for structured small groups have been obtained (Johnson and Johnson 1999; Johnson et al. 2000). Therefore, the third hypothesis was that structured small-group learning methods would yield higher effect sizes than unstructured ones on transfer.

Method

The population consists of studies investigating the effect of small-group learning on academic-related transfer achievement. The databases used for searching studies included *Educational Resources Information Center* (*ERIC*), *PsycINFO*, *Digital Dissertations Database*, search engines such as *Google Scholar*, and the reference lists of the identified articles. The keywords used for search included "collaborative learning" OR "cooperative learning" OR "dyads" OR "group-based learning" AND "transfer". Because "transfer" is a term with many potential conceptualizations (cf. Barnett and Ceci 2002; Lobato 2006), studies had to explicitly use the term "transfer" in order to be included in the meta-analysis. This criterion is critical in our study because it helped avoid potentially subjective judgments about what types of assessments would count as transfer assessments. More importantly, it helped differentiate this study from previous reviews of cooperative learning. For example, Slavin (1995) included studies in his review that involved performance on standardized tests after multiple weeks of cooperative learning or control treatment. Because standardized tests are different from experimenter-made tests—they tend to be less targeted to the intervention and may present questions in ways that appear different on the surface from what students encountered during initial learning—they could be seen as including measures of transfer. However, without providing a separate score for transfer items (versus factual recall items, for example), the overall standardized test score may provide a metric that is unrepresentative of effects specifically on transfer (i.e., construct irrelevant variance). Thus, if such studies were all included in this meta-analysis, they would risk blurring the potential distinction between well-established outcomes of cooperative learning versus specific effects on transfer.

However, requiring the use of the term "transfer" has potential limitations mostly associated with missing studies that examine transfer performance but refer to it with a different term. We suspect it would be rare for studies differentiating learning performance from transfer performance of groups versus individuals not to use the term "transfer" *and* simultaneously to obtain dramatically different results from those studies that do use the term. For example, Cohen et al. (1999) reported results that might fall into this "missed" category. Their results were similar to some of the previously reviewed studies that contrasted factual recall with transfer. Collaborative learning was not associated with significant advantages on factual recall, but it was associated with significant advantages on (transfer) measures of conceptual understanding (Cohen et al. 1999). Another possible concern is that relying on other's use of the term "transfer" puts too much faith in their operationalization of the term. Studies were excluded from the analysis if their operationalization of the transfer test did not fit the notion of adapting or applying prior knowledge to a novel context. For example, if a study used the same assessment at pretest and posttest (e.g., Souvignier and Kronenberger 2007), then it was excluded from the study because the posttest would lack novelty (i.e., the retrieval cues provided by the same test could remove many of the traditional barriers to transfer that tend to distinguish it from learning).

In addition to the explicit use of the term "transfer", to qualify for inclusion in the meta-analysis, studies had to include: (a) a research design that compared at least one condition using small-group learning to at least one condition using individual learning, (b) an assessment, taken individually after the learning phase, specifically designed to measure transfer, and (c) data on academic-related transfer performance that supported calculation of effect sizes comparable across studies (e.g., means and standard deviations of students' transfer performance scores, or *t*-statistics). Academic-related content is the main interest of this study, and it was used here to refer to school subject matter, including mathematics, science, biology, ecology, geology, astronomy, music, and vocabulary, as well as related activities including logic problems. The meta-analysis did not include studies of transfer in the work place.

The first author and three graduate students were engaged in study identification under supervision of the second author, who has expertise in collaborative learning and transfer. One hundred and thirty abstracts were reviewed after the initial search with the above-mentioned keywords in the listed databases. Many of the 130 studies were excluded from our sample pool after reading the abstracts because they did not meet the inclusion criteria. The search was further refined by reading the remaining articles in detail. Fifteen articles meeting the criteria were identified through the process. The references of the selected articles were also screened for finding additional studies, and nine articles were identified through the reference lists or the relevant authors. A total of 24 empirical published and unpublished studies, which were reported in 20 journal articles, 2 dissertations, and 2 conference papers, were identified. Nearly all of the studies (19 of 24) were published since 2000, and 11 have appeared since 2008.

Publication Bias and Power Analysis

Ideally, all studies that meet the inclusion criteria of a systematic review are located and included in a meta-analysis. However, several issues make it rarely possible to reach this ideal (Borenstein et al. 2009). For example, studies with significant results are more likely to be published than studies reporting non-statistically significant results. Unpublished studies and the literature that is not controlled by commercial publishers are less likely to be located and included in a meta-analysis. Other factors such as language bias (English-language databases are more likely to be searched), availability bias (easily accessible studies are more likely to be selected), and citation bias (studies with statically significant results are more likely to be cited by other studies and thus easier to be identified) may lead to a biased overall effect size.

To limit potential publication bias, a thorough search was conducted by including multiple databases with both published and unpublished studies, and also by contacting authors when encountering studies reporting insufficient information. For example, we contacted Kramarski to obtain standard deviations in their study (Kramarski and Mevarech 2003). We also examined a funnel plot (Light and Pillemer 1984) to display the relationship between effect size and study size to illustrate potential evidence of publication bias. Figure 1 illustrates that larger studies are distributed symmetrically on the top and in the middle, and a few small-scale studies are missing near the bottom. This indicates a gap on the left near the bottom where the small-scale studies with low or negative effect sizes could have been if they could be located. In practice, studies with small sample size and nonsignificant results are less likely to be reported. On the other hand, these studies may be less influential on the meta-analytic results because they tend to provide a small weight in the weighted average effect size computation.

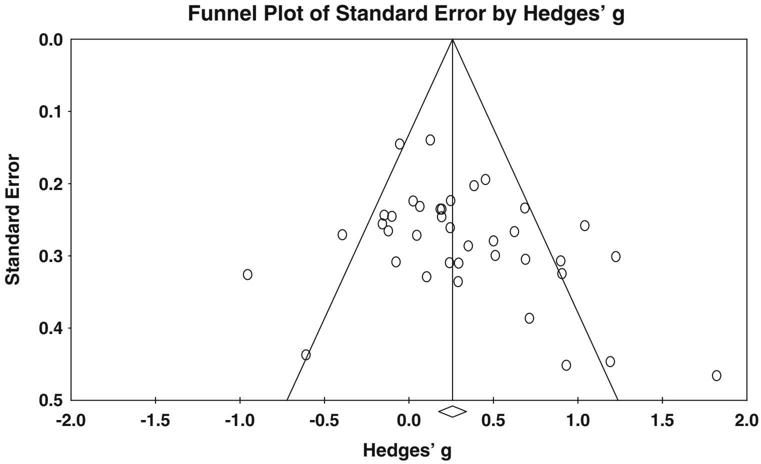

Fig. 1 Funnel plot of standard error by Hedges' *g* effect sizes

Another issue of meta-analysis, that is similar to primary studies, is having sufficient statistical power to detect a main effect. We performed an a priori power analysis and found that to have power of 90 % for an effect size of 0.20, 30 cases would be needed. Thus, this meta-analysis with 38 effect sizes should possess sufficient power to detect main effects.

Codes and Coding Procedures

Table 1 summarizes the key study variables, which include the characteristics of report (e.g., source of the documentation), study setting (e.g., location of the study), participant characteristics (e.g., grade level of the participants), intervention characteristics (e.g., types of small-group learning), and outcome measure (e.g., type of outcome measure including standardized tests, school tasks, and puzzles). Note that among the coded moderators, the structured and unstructured small-group learning intervention characteristic formed the theory-based moderator of interest in this meta-analysis. When guidance of group interaction was provided, such as with scripts, assigned roles, group goals, and individual accountability, or group-skills training, this kind of small-group learning was categorized as structured. In contrast, when group members were simply asked to work together, this was categorized as unstructured small-group learning. Comparing amongst different structuring methods is not the purpose of this study. Therefore, the dichotomous category of structured versus unstructured small-group learning was used.

The first author and the same three graduate assistants who searched studies also served as coders. Coding consistency and accuracy were assured by: (a) training coders on several articles, (b) independent coding of each article by at least two members of the coding team with disagreements resolved by discussion, (c) re-examination of the codes by the first author, and (d) supervision of the process by the second author.

Effect Size Extraction

The Hedges' *g* effect sizes served as outcomes in the current study because Cohen's *d* tends to overestimate the value of standardized mean difference when sample sizes are small (Borenstein et al, 2009). A positive *g* effect size would indicate that small-group learning

Table 1 Study variables

Variable	Description	Code
Grade level	Participants' grade level	1=1st grade…12=12th grade
		13=Freshman of college…
		16=Senior of college
		17=other
Instructional method	The type of teaching method implemented in the primary study	1=Traditional method (lesson and practice)
		2=Discovery learning (including problem solving)
		3=Other
Type of small-group learning	Whether or not the group interaction was structured.	0=Unstructured group work
		1=Structured group work (e.g., script, assigned role, interdependence & individual accountability, group goal, individual accountability, etc.)
Content area	The content area of students' academic performance investigated	1=Reading
		2=Other language arts
		3=Math
		4=Science
		5=Social Studies
		6=Foreign language
		7=Other
		8=Not a subject matter test
Type of outcome measure	The type of outcome measure used in the primary study	1=School task
		2=Standardized achievement test
		3=Researcher arbitrary puzzle
		4=Multiple types of measures
		5=Other

conditions outperformed the individual learning conditions on transfer tasks while a negative effect size would indicate an advantage for individuals.

When multiple effect sizes were reported in the same study (e.g., Hudgins 1960; Kirschner et al. 2011; Kramarski and Mevarech 2003; Krause and Stark 2010), they were treated as independent effect sizes if they were from independent samples. Effect sizes can also come from studies comparing multiple types of treatments against a common control condition (multiple-treatment studies), or some studies may have one treatment and one control but include multiple measures on the same participants (multiple-endpoint studies). If the effect sizes were from dependent groups and were still aggregated with other independent ones, it causes bias and as a result, threatens the validity of the meta-analysis (e.g., Wood 2008). To handle this potential problem, the authors decided to select one representing effect size if the primary study reported multiple effect sizes from dependent samples. This approach avoids complications of moderator analysis when averaging multiple effect sizes with different sample sizes. It is also more feasible than the statistical adjustment approach which requires additional statistics that primary studies may not report (Borenstein et al. 2009). A weakness of selecting a representative effect size is that determining which effect size to be included is subjective, thus it requires clear selection strategies. When multiple dependent effect sizes were reported, we first examined the quality of each dependent effect size and selected the one that best represents the purpose of the meta-analysis. For example, Barron's (2000) study included multiple-endpoints. There were three measures to assess students' different problem solving skills on a transfer task, and it resulted in three effect sizes. The first two measures, general planning questions and subproblem planning questions, were designed to assess students' thinking process during problem solving. They were not similar to the performance measures in other studies. Therefore, the third effect size, assessing accuracy of participants' final solutions, was selected. Moreno's (2009) study provides an example of multiple-treatments. She investigated the effects of three different instructional methods for college students learning about botany, including individual learning, a jigsaw method, and a cooperative learning method. Among the three approaches, the individual condition served as the control in this meta-analysis, and the jigsaw method and cooperative learning method were the treatments. Thus, two effect sizes were yielded from comparing two treatments to one control condition. To select the representing effect size, the quality of the research design was assessed. Participants in the jigsaw condition were randomly assigned to learn one of the three modules of the material and taught what they learned to their peers. However, unlike typical jigsaw requirements (Aronson 2002), there were no expert-group meetings before students shared their specialized knowledge with their peers. Moreover, because no other study included in the meta-analysis used the jigsaw method, the effect size calculated from this condition was excluded in favor of the one from the "cooperative learning" treatment. Through this process, we identified 38 independent effect sizes from the 24 studies. The total number of research participants across studies was 3,106. Table 2 summarizes the studies with key study features that were included in the current meta-analysis.

Data Analysis

The analytical framework selected for the current meta-analysis is the random-effects model. As opposed to the fixed-effects model, the assumption of the random-effects model is that the true effect size is not exactly the same in all studies. In this meta-analysis, the extracted effect sizes were from studies that have different characteristics of samples, research design and procedures, each of which could contribute to heterogeneity among the population effects. Thus, we considered the random-effects model to be appropriate (Hedges and Vevea 1998). To obtain a precise estimate for the overall mean effect size of the population under the random-effects model, weighting was used. The weight of a given study is the inverse of that study's variance, which is the combination of the within-study variance (i.e., sampling error) and the between-studies variance (T^2, the estimated variance of the effect sizes across the population of studies) (Borenstein et al. 2009).

A histogram and a forest plot with 95 % confidence interval for each independent effect size was created to detect patterns in the magnitude of the effect sizes. A funnel plot (Light and Pillemer 1984) was created to detect the existence of publication bias. The weighted average of all independent effect sizes was computed to quantify the central tendency among the effect sizes.

To investigate the variation of the effect sizes across studies three relevant statistics from the random-effects model of meta-analysis were examined: (a) the chi-square test of homogeneity of effect sizes (Q test, Hedges and Olkin 1985), (b) the tau-squared statistic (T^2), and (c) the I^2 statistic (Higgins et al. 2003). If statistically significant variation exists among effect sizes, as expected, inferential analyses would be conducted to identify moderators to explain the variation.

Table 2 Coded study characteristics

Study	ID	Type	Education Level	Group Type	Group Size	Transfer Task	N_G	N_I	d	Var_d	g	Var_g
Barab et al. (2009)	Ex	J	UG	Unstructured	2	Multiple choices	12	15	0.41	0.15	0.40	0.14
	1	J	UG	Unstructured	2	Open-ended performance-based transfer questions	12[a]	15[a]	0.74	0.16	0.72	0.15
Barron (2000)	Ex	J	MS	Unstructured	3	Problem analogous to Jasper (general planning)	48	48	0.26	0.04	0.26	0.04
	Ex	J	MS	Unstructured	3	Problem analogous to Jasper (subproblem planning)	48	48	0.22	0.04	0.21	0.04
	2	J	MS	Unstructured	3	Problem analogous to Jasper (solutions)	48	48	0.40	0.04	0.39	0.04
Brand et al. (2003)	3	J	UG	Unstructured	2	Tower of Hanoi	26	26	0.05	0.08	0.05	0.07
	4	J	UG	Unstructured	2	Tower of Hanoi	25	30	−0.12	0.07	−0.12	0.07
	Ex	J	UG	Unstructured	2	Missionary and Cannibal problem, and Katona Card problem	26	26	−0.13	0.08	−0.13	0.08
	Ex	J	UG	Unstructured	2	Missionary and Cannibal problem, and Katona Card problem	25	30	0.07	0.08	0.07	0.07
Brodbeck and Greitemeyer (2000)	5	J	UG	Structured (reward)	3	Rule-induction card tasks	22	22	0.53	0.09	0.52	0.09
Corliss (2005)	6	D	UG	Unstructured	2	Near transfer: problem analogous to the learning task	36	34	0.19	0.06	0.19	0.06
	7	D	UG	Unstructured	2	Problem analogous to the learning task	39	39	0.26	0.05	0.25	0.05
	8	D	UG	Unstructured	2	Problem analogous to the learning task	38	39	0.03	0.05	0.03	0.05
	9	D	UG	Unstructured	2	Problem analogous to the learning task	36	36	0.07	0.06	0.07	0.05
	Ex	D	UG	Unstructured	2	Far transfer task	36	35	0.04	0.05	0.04	0.06
	Ex	D	UG	Unstructured	2	Far transfer task	39	39	−0.43	0.05	−0.42	0.05
	Ex	D	UG	Unstructured	2	Far transfer task	38	39	0.01	0.05	0.01	0.05
	Ex	D	UG	Unstructured	2	Far transfer task	35	35	−0.30	0.06	−0.29	0.06
Cuneo (2007)	10	D	UG	Structured (Individual Accountability)	2	Near transfer: algebra problems	34	36	0.20	0.06	0.20	0.06
Hudgins (1960)	11	J	E	Unstructured	4	Stanford achievement test	32	32	−0.10	0.06	−0.09	0.06
	12	J	E	Unstructured	4	Stanford achievement test	32	32	0.20	0.06	0.20	0.06

Table 2 (continued)

Study	ID	Type	Education Level	Group Type	Group Size	Transfer Task	N_G	N_I	d	Var_d	g	Var_g
Kapur (2010)	13	J	MS	Unstructured	2 or 3	Higher-order analysis & application item	37	38	0.70	0.06	0.69	0.06
Kirschner et al. (2009)	14	J	HS	Structured (task interdependence)	3	Three transfer problems structurally different from learning tasks	18	16	0.30	0.12	0.30	0.11
Kirschner et al. (2011)	15	J	HS	Structured (task interdependence)	3	Four transfer test using same underlying theory with learning task	34	32	1.06	0.07	1.05	0.07
	16	J	HS	Structured (task interdependence)	3	Four transfer test using same underlying theory with learning task	33	32	−0.14	0.06	−0.14	0.06
Kramarski and Mevarech (2003)	17	J	MS	Structured (metacognitive training+scripts)	4	Transfer verbal description into graphic representations	105	95	0.13	0.02	0.13	0.02
	18	J	MS	Structured (individual accountability)	4	Transfer verbal description into graphic representations	91	93	−0.05	0.02	−0.05	0.02
Krause and Stark (2010)	19	J	UG	Unstructured	2	Three transfer tasks demanded declarative, procedural and conditional knowledge	50	17	0.51	0.08	0.51	0.08
	20	J	UG	Unstructured	2	Three transfer tasks demanded declarative, procedural and conditional knowledge	52	18	−0.39	0.08	−0.39	0.07
Lambiotte et al. (1987)	Ex	J	UG	Structured (scripts)	2	Transfer study/test-taking strategies in different content	18	17	0.55	0.12	0.52	0.11
	21	J	UG	Structured (scripts)	2	Transfer study/test-taking strategies in different content	18	17	0.11	0.12	0.11	0.11
Laughlin et al. (2008)	22	J	UG	Unstructured	3	Find out the assigned digit to each letter	105	35	0.46	0.04	0.46	0.04
	Ex	J	UG	Unstructured	3	Find out the assigned digit to each letter	105	35	0.41	0.04	0.41	0.04
	Ex	J	UG	Unstructured	3	Find out the assigned digit to each letter	105	35	0.32	0.04	0.32	0.04
McDonald et al. (1985)	Ex	J	UG	Structured (scripts)	2	Transfer study/test-taking strategies in different content	20	20	0.70	0.11	0.68	0.10
	23	J	UG	Structured (scripts)	2	Transfer study/test-taking strategies in different content	20	20	−0.07	0.10	−0.07	0.10
	24	J	UG	Structured (scripts)	2	Transfer study/test-taking strategies in different content	30	27	0.64	0.07	0.63	0.07
	Ex	J	UG	Structured (scripts)	2	Transfer study/test-taking strategies in different content	30	30	0.75	0.07	0.74	0.07
Mondoux et al. (2004)	25	C	UG	Unstructured	2	Analyze 5 situations and answer 3 questions measured transfer	28	23	1.25	0.10	1.23	0.09

Table 2 (continued)

Study	ID	Type	Education Level	Group Type	Group Size	Transfer Task	N_G	N_I	d	Var_d	g	Var_g
Moreno (2009)	Ex	J	UG	Structured (jigsaw)	3	A problem-solving test	30	27	−0.66	0.07	−0.65	0.07
	26	J	UG	Structured (assigned roles)	3	A problem-solving test	30	27	0.25	0.07	0.25	0.07
O'Donnell et al. (1985)	27	J	UG	Unstructured	2	Instruction writing task that requires to transfer strategies but not content knowledge (communicativeness)	9	18	1.89	0.23	1.83	0.22
	Ex	J	UG	Unstructured	2	Instruction writing task that requires to transfer strategies but not content knowledge (completeness)	9	18	0.67	0.23	0.65	0.17
O'Donnell et al. (1986)	28	J	UG	Unstructured	2	Instruction writing task that requires to transfer strategies but not content knowledge (communicativeness)	24	21	0.92	0.10	0.90	0.10
	Ex	J	UG	Unstructured	2	Instruction writing task that requires to transfer strategies but not content knowledge (completeness)	24	21	−0.01	0.09	−0.01	0.09
	29	J	UG	Unstructured	2	Instruction writing task that requires to transfer strategies but not content knowledge (communicativeness)	24	20	0.71	0.10	0.70	0.09
	Ex	J	UG	Unstructured	2	Instruction writing task that requires to transfer strategies but not content knowledge (completeness)	24	20	0.41	0.09	0.40	0.09
Olivera and Straus (2004)	30	J	UG	Unstructured	3	Brainteaser puzzles (identify word/phrase depicted in figure)	18	35	0.36	0.09	0.36	0.08
Rebetez et al. (2010)	31	J	UG	Unstructured	2	Questions have same principles but applied to other phenomena	20	20	0.25	0.10	0.25	0.10
	32	J	UG	Unstructured	2	Questions have same principles but applied to other phenomena	20	20	−0.97	0.11	−0.95	0.11
	33	J	UG	Unstructured	2	Questions have same principles but applied to other phenomena	20	20	0.93	0.11	0.91	0.11
	34	J	UG	Unstructured	2	Questions have same principles but applied to other phenomena	20	20	0.31	0.10	0.30	0.10
Sears et al. (2011)	35	C	UG	Unstructured	2	Multiple-choice test of definitions of new words	11	11	1.25	0.22	1.20	0.20
Sears and Pai (2012)	36	J	UG	Unstructured	2	Dictation of sixth musical phrase	10	10	0.98	0.22	0.94	0.21

Table 2 (continued)

Study	ID	Type	Education Level	Group Type	Group Size	Transfer Task	N_G	N_I	d	Var_d	g	Var_g
	37	J	UG	Structured (group goal+ individual accountability)	2	Dictation of sixth musical phrase	10	10	−0.63	0.21	−0.60	0.19
	Ex	J	UG	Unstructured	2	Dictation of sixth musical phrase	10	10	0.78	0.22	0.74	0.20
	Ex	J	UG	Unstructured	2	Dictation of sixth musical phrase	10	10	0.65	0.21	0.62	0.19
Wirkala and Kuhn (2011)	38	J	MS	Structured (scripts+assigned roles)	3	Topic 1 (groupthink): application assessment	29	30	−0.15	0.07	−0.15	0.07
	Ex	J	MS	Structured (scripts+assigned roles)	3	Topic 1 (groupthink): application assessment	29	31	0.85	0.07	0.84	0.07
	Ex	J	MS	Structured (scripts+assigned roles)	3	Topic 2 (memory): application assessment	30	31	−0.18	0.07	−0.18	0.06
	Ex	J	MS	Structured (scripts+assigned roles)	3	Topic 2 (memory): application assessment	30	29	0.58	0.07	0.57	0.07

Note. The positive effect size indicates small-group learning condition has higher performance on transfer test than individual learning

ID the identification number of selected independent effect size (*Ex* excluded effect size due to dependency), *Type* publication type (*C* conference proceeding, *D* dissertation, *J* journal), *E* elementary school students, *MS* middle school students, *HS* high school students, *UG* undergraduate students, N_G sample size for groups, N_I sample size for individuals, *d* Cohen's *d* indicating the standardized mean difference (group average−individual average), *g* Hedges' measure of effect size, Var_d variance of *d*, Var_g variance of *g*

[a] Estimated number

As previously noted, the primary moderator of interest was the type of small-group learning (structured versus unstructured).

Results

Distribution of the 38 Hedges' *g* effect sizes is presented in Figure 2. The graph shows a roughly normal distribution with many of the effect sizes clustering between 0.0 and 1.0. Figure 3 shows the distribution of the 38 independent effect sizes with their 95 % confidence intervals. The mid-point of the diamond represents the point estimate of each effect size, and the width of the line shows the 95 % chance the true effect will lie within the range. The plot indicates the variation of 38 effect sizes with relatively large confidence intervals. The weighted overall average of the 38 independent effect sizes for the difference between small-group and individual learning on transfer performance was 0.30 (SE=0.07, 95 % CI=0.16 to 0.44). It indicates that learners who studied in groups outperformed their peers on transfer tasks who studied alone by about one-third of a standard deviation, on average.

Heterogeneity of Effect Sizes

The chi-square test of homogeneity of effect sizes indicated that effect sizes varied significantly across studies, $Q\,(37)=98.88$, $p<0.01$. The tau-squared (T^2) refers to the estimation of the variance of effect sizes, $T^2=0.11$, and it revealed sizable variation in parameter effect sizes. The I^2 statistic was 62.58 %, which indicates that a high proportion of the between-effect size variance reflects real differences in effect sizes. Thus, it is reasonable to speculate that differences between studies might explain the variance and a moderator analysis is appropriate.

For the moderator analysis, there were 12 independent effect sizes derived from 10 studies which used structure to enhance small-group learning interaction. Of the 12, nine effect sizes were from the conditions under motivational structures that provided incentives to encourage help among group members, including group rewards based on individual performance (Brodbeck and Greitemeyer 2000; Sears and Pai 2012), individual accountability (Cuneo 2007; Kramarski and Mevarech 2003; Moreno 2009), and task specialization, where each

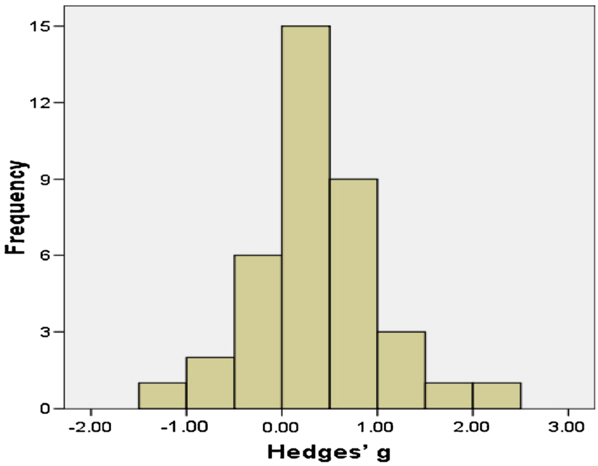

Fig. 2 Distribution of 38 Hedges' *g* effect sizes

participant received a portion of the total lesson (Kirschner et al. 2009; Kirschner et al. 2011). Two effect sizes were from the conditions that used assigned roles to direct group interaction (Lambiotte et al. 1987; McDonald et al. 1985). One effect size used multiple structuring methods, including assigning a team leader and a script (Wirkala and Kuhn 2011).

Of the 26 unstructured effect sizes from 15 studies, five studies which provided seven effect sizes added minor supports for interaction that went beyond simply asking students to work together. For example, Corliss (2005) applied reflective questions at three different times during the research session to prompt students' metacognitive thinking. She also included a transfer survey to prompt students to consider the usefulness of the problem-solving process in the future. Students could choose to discuss the prompting questions together (or not) in the group condition. Thus, students could simply work on their own prompting questions without collaborating with their peers. In other words, these were minor supports because they did not script the interaction process. Three other studies used similar supports and were also categorized as unstructured. They used metacognitive questions about identifying problems, establishing goals, forming and testing hypotheses, and setting new plans of action to prompt students' thinking (Barab et al. 2009; Brand et al. 2003; Hudgins 1960). In these three studies, aside from requiring consensus, groups had considerable freedom to determine how they would engage in the task. In other words, there was little in these task requirements that would prevent social loafing, dominance of high status students, or acquiescence of others. The fifth study, by Krause and Stark (2010), provided minimal support for interaction in the form of feedback. In their study, students took multiple-choice tests with immediate feedback about why each answer was correct or incorrect. This study was categorized as unstructured because the feedback was not an external structure designed to shape group interactions.

Figure 3 also depicts the variation of effect sizes by the nature of small group learning with outlined diamonds representing unstructured small-group learning and filled diamonds representing structured. The weighted average effect size of the unstructured small-group learning studies was 0.36 (SE=0.09, 95 % CI=0.17 to 0.54, n=26). It indicates that students' transfer performance in unstructured small-group learning conditions was about 0.36 standard deviations higher than that in the individual learning conditions. It was not significantly different from students' transfer performance in structured small-group learning conditions of 0.20 (SE=0.11, 95 % CI=−0.01 to 0.41, n=12), Q (1)=1.251, p=0.263. Although these results are not significantly different from each other, it is important to note that only the unstructured small-groups showed significant advantages over their respective individual-study conditions, as indicated by the confidence intervals.

Discussion

The estimated overall effect size of small-group learning on students' transfer performance was 0.30, suggesting that small-group learning had a superior impact on students' performance on transfer tasks compared to individual learning. This finding is consistent with reviews of effects of structured cooperative learning (Johnson et al. 2000; Slavin 1995, 1996). The current meta-analysis adds to these findings by showing that small-group learning can have similar positive benefits for transfer as for learning more generally. In addition, it suggests that learning in small groups may naturally support transfer, even without the use of external structures, such as scripts, roles, or rewards.

Moderator analysis indicated that the effect of small group learning on students' transfer performance was not significantly different for structured versus unstructured conditions; however, it also suggested that the unstructured studies were largely responsible for small-groups

Study	ID	Hedges' g
O'Donnell, Dansereau, Rocklin, Lambiotte, Hythecker & Larson (1985)	27	1.83
Mondoux, Auderset,& Dillenbourg (2004)	25	1.23
Sears, Packer, Hoffa, Min, Sasser, & Simison(2011)	35	1.20
Kirschner, Paas, Kirschner,& Janssen (2011)	15	1.05
Sears & Pai (2012)	36	0.94
Rebetez, Bétrancourt , Sangin, & Dillenbourg (2010)	33	0.91
O'Donnell, Larson, Dansereau, & Rocklin (1986)	28	0.90
Barab, Scott, Siyahhan, Goldstone, Ingram-Goble, Zuiker, & Warren (2009)	1	0.72
O'Donnell, Larson, Dansereau, & Rocklin (1986)	29	0.70
Kapur (2010)	13	0.69
McDonald, Larson, Dansereau, & Spurlin (1985)	24	0.63
Brodbeck & Greitemeyer (2000)	5	0.52
Krause & Stark (2010)	19	0.51
Laughlin, Carey & Kerr (2008)	22	0.46
Barron (2000)	2	0.39
Olivera & Straus (2004)	30	0.36
Rebetez, Bétrancourt , Sangin, & Dillenbourg (2010)	34	0.30
Kirschner, Paas, & Kirschner (2009)	14	0.30
Corliss (2005)	7	0.25
Moreno (2009)	26	0.25
Rebetez, Bétrancourt , Sangin, & Dillenbourg (2010)	31	0.25
Cuneo (2007)	10	0.20
Hudgins (1960)	12	0.20
Corliss (2005)	6	0.19
Kramarski & Mevarech (2003)	17	0.13
Lambiotte, Dansereau, Rocklin, Fletcher, Hythecker, Larson, & O'Donnell (1987)	21	0.11
Corliss (2005)	9	0.07
Brand, Reimer, & Opwis (2003)	3	0.05
Corliss (2005)	8	0.03
Kramarski & Mevarech (2003)	18	-0.05
McDonald, Larson, Dansereau, & Spurlin (1985)	23	-0.07
Hudgins (1960)	11	-0.09
Brand, Reimer, & Opwis (2003)	4	-0.12
Kirschner, Paas, Kirschner,& Janssen (2011)	16	-0.14
Wirkala & Kuhn (2011)	38	-0.15
Krause & Stark (2010)	20	-0.39
Sears & Pai (2012)	37	-0.60
Rebetez, Bétrancourt , Sangin,& Dillenbourg (2010)	32	-0.95

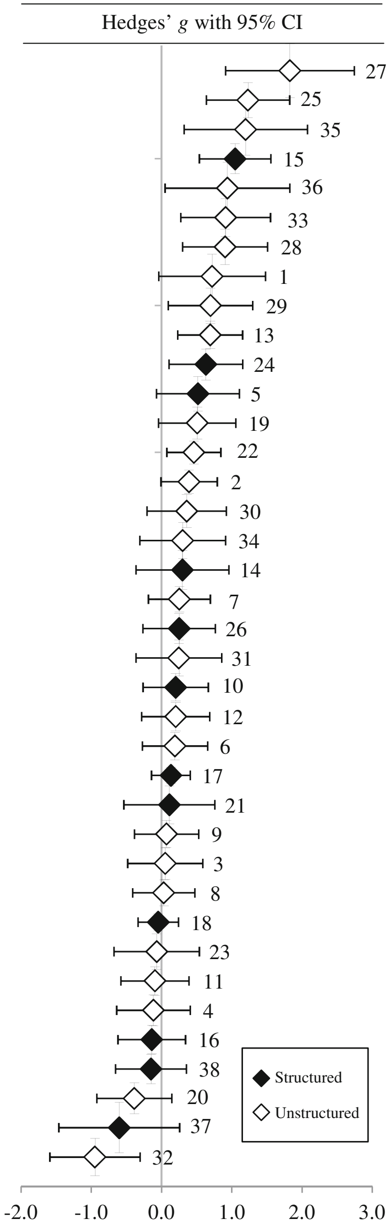

Fig. 3 Hedges' *g* effect sizes with unstructured/structured small-group learning

surpassing individuals on transfer. Given previous reviews highlighting benefits of structuring cooperative learning, this result is surprising (cf., Johnson et al. 2000; O'Donnell 2006; Slavin 1995). One possible explanation for this result is that the definition of structure was broader in this meta-analysis than in previous reviews, which tended to focus on cooperative learning with positive interdependence (Johnson et al. 2000; Slavin 1995). More studies with both positive interdependence and measures of transfer performance are needed to examine this possibility.

Another possible explanation is that potential benefits of structure for transfer are affected by other factors, such as task complexity, group size, or student age. For example, if a group is relatively large and a task is complex, then providing structure may help avert common problems that tend to become worse as group size increases, such as social loafing (Latané et al. 1979). A similar argument regarding cognitive mechanisms could also apply. For instance, as tasks increase in complexity and groups increase in size, at some point working memory may become overloaded (Kirschner et al. 2009). Under these conditions, adding structure might be particularly beneficial, assuming it serves to reduce cognitive load to manageable levels and thereby facilitates other processes by which deeper understanding can occur, such as giving and receiving explanations (Webb 1982). In the current study, only two studies included four-member groups; the rest consisted of dyads or triads. Clearly, more primary studies with larger group sizes and tasks of different complexity levels are needed to further delineate when and how structure might support transfer.

If future reviews continue to find advantages of unstructured (vs. structured) small-group learning for transfer, studies comparing the interaction processes that occur in structured versus unstructured versions of the same task could help differentiate processes primarily associated with transfer versus those primarily associated with initial acquisition. This could provide important new insights both for the field of small-group learning and for the field of transfer. However, given that Slavin's (1995) review highlighting the importance of motivational structures for successful cooperative learning involved many studies with prolonged durations and standardized tests as dependent measures—tests which presumably require greater transfer than many immediately administered experimenter-made tests of learning—there is reason to suspect that some other factor, such as group size or task complexity, is responsible for this unexpected result in favor of unstructured small groups.

Regardless of the eventual answer to this puzzle, these results highlight an important point about small-group learning. Despite the many forms that group-work can take, a fairly consistent outcome was benefits to transfer. This suggests that the overlap between cognitive processes associated with transfer and with collaboration is fairly robust. It also highlights the importance of including measures of transfer in studies of small-group learning. Without measures of transfer, a number of studies in this meta-analysis would have shown no differences between individual and small-group conditions (e.g., Kirschner et al. 2009; Moreno 2009; Rebetez et al. 2010; Sears et al. 2011).

The 24 studies included in this meta-analysis used different tasks. For example, some studies included other treatment variables, such as metacognitive stimulation or training (e.g., Brand et al. 2003; Corliss 2005; Kramarski and Mevarech 2003), assigning roles of summarizer and listener (McDonald et al. 1985), using different learning-material formats (Rebetez et al. 2010) or different instructional methods (Wirkala and Kuhn 2011). It remains to be seen if the significant variability in effect sizes of these widely varied tasks can be captured by one or two simple moderators, such as task type or complexity.

Limitations and Future Directions

While the conclusion that small-group learning benefits transfer on school-related tasks is theoretically and now empirically well-supported, caution is warranted in interpreting the results of the moderator analysis comparing structured versus unstructured small groups. A primary limitation is that relatively few of the included studies involved young students or relatively large groups, for whom greater scaffolding may be needed for productive group interaction. In addition, for the included studies, the operationalization of various structures often failed to ensure a strong degree of positive interdependence. For example, one of the effective methods of structuring cooperative learning is to provide group rewards based on an aggregate of group members' individual performances, which is known to foster positive interdependence (Slavin 1996; Johnson and Johnson 1999). In Slavin's review (1995, 1996), studies that only used group goals without individual accountability or provided no group goals found little positive effect. Among the 24 primary studies in the current meta-analysis, five studies applied the concept of group goals and/or individual accountability. However, only two studies used individual accountability to the group goal (Brodbeck and Greitemeyer 2000; Sears and Pai 2012). Other studies either applied only group goals or individual accountability or used separate tasks where one had a group goal (e.g., a group project) and the other had individual accountability (an individual test) (Cuneo 2007; Kramarski and Mevarech 2003; Moreno 2009). This latter structure deserves further comment because on the surface it appears to have group goals plus individual accountability. However, closer inspection reveals that the individual accountability is not to the group outcome. The individuals were only accountable to themselves on the test. Furthermore, the group project did not trace individual contributions. We suspect this is a common misunderstanding of group goals plus individual accountability and suggest that such methods should not be expected to show the consistently positive results associated with methods that use individual accountability to the group goal because they may not promote positive interdependence (e.g., Johnson and Johnson 1999, 2009; Slavin 1995, 1996).

A related limitation was the small sample of studies that used each type of small-group structure, such as group rewards versus scripts and roles. This prevented meaningful analyses of which type(s) of structure might be most effective for promoting transfer, thereby limiting the moderator analysis to the dichotomous category of structured versus unstructured. This observation provides a strong suggestion for future research, especially given the differences noted between different types of structures in previous reviews (Johnson and Johnson 2009; Slavin 1996). More studies that implement different types of group structures and test their effects on transfer are needed.

One way that future meta-analyses could increase the sample of studies reviewed is by broadening the inclusion criteria. As mentioned earlier, this study required the use of the term "transfer" in order: (a) to help avoid potentially subjective judgments regarding which studies lacking the term should be included and (b) to distinguish potential effects of small groups on transfer from well-established effects on learning. While the explicit use of the term "transfer" ensures the validity of this meta-analysis[1], this necessarily excludes studies that use assessments of "higher-order learning" or "conceptual understanding" but fail to use the term "transfer." It is worth noting that conceivably, some of these studies might offer even better measures of transfer than those in the studies meeting the inclusion criteria. This possibility highlights the importance of future work further delineating whether and when transfer and conceptual understanding (or related constructs) are practically distinguishable in collaborative contexts. Empirically and theoretically, there are reasons to expect that the benefits seen for transfer in the current study would extend to other similar measures of conceptual understanding of complex material even when they may not apply to immediate factual recall (e.g., Cohen 1994; Cohen et al. 1999; Kirschner et al. 2009; Sears and Reagin 2013).

While the limited number of studies with proven positive-interdependence structures may explain why structured small-groups showed no greater benefit for transfer than unstructured, they do not explain why unstructured performed so well (average effect size of 0.36). If

[1] Although transfer can have different conceptualizations and operationalizations (e.g., near vs. far; direct application vs. dynamic transfer, Schwartz et al. 2008), there is broad consensus that transfer involves the generalization of prior knowledge, and this is the broad sense of the term that is meant here.

unstructured group interaction works as well as structured for promoting transfer, teachers would have additional information about when they could find educational benefits from unstructured small-group learning and could plan accordingly. The significant variability in effect sizes in this meta-analysis were not explained by the theory-based (structured vs. unstructured) moderator. Additional research is needed to explain this variance to provide more information about under what conditions small-group learning can promote transfer. One factor that warrants further investigation is task-effects, especially task complexity (e.g., Cohen 1994; Kirschner et al. 2009; Laughlin et al. 2008; Sears and Reagin 2013). With many of the studies included in this meta-analysis conducted within the last 5 to 10 years, the empirical relationship between small-group learning and transfer appears to be a burgeoning research area that has plenty of questions in need of answers and the potential for synergistic constructs and methods.

In summary, this meta-analysis offered potential insights for both fields: small-group learning and transfer. By investigating the effect of small-group learning on transfer performance, it revealed hidden benefits of small-group learning that traditional measures overlook. In the field of transfer, it is widely recognized that researchers have often struggled to promote transfer reliably (Detterman 1993; Lobato 2006). This study suggested one relatively simple method for increasing the occurrence of transfer.

Acknowledgements The authors would like to thank Dr. Helen Patrick, Dr. Victoria Walker, Dr. Scott Schaffer, and anonymous reviewers for feedback on earlier versions of this work.

References

References marked with an asterisk indicate studies included in the meta-analysis

Andersson, J., & Rönnberg, J. (1995). Recall suffers from collaboration: joint recall effects of friendship and task complexity. *Applied Cognitive Psychology, 9*(3), 199–211.

Aronson, E. N. (2002). The jigsaw classroom. In E. N. Aronson (Ed.), *Improving academic achievement: impact of psychological factors on education* (pp. 215–219). San Diego: Academic Press.

Aronson, E., Blaney, N., Stephan, C., Sikes, J., & Snapp, M. (1978). *The jigsaw classroom*. Beverly Hills: Sage.

Ashman, A. F., & Gillies, R. M. (1997). Children's cooperative behavior and interaction in trained and untrained work groups in regular classrooms. *J Sch Psychol, 35*(3), 261–279.

*Barab, S. A., Scott, B., Siyahhan, S., Goldstone, R., Ingram-Goble, A., Zuiker, S. J., & Warren, S. (2009). Transformational play as a curricular scaffold: using video games to support science education. *Journal of Science Education and Technology*, *18*(4), 305–320.

Barnett, S. M., & Ceci, S. (2002). When and where do we apply what we learn? *Psychol Bull, 128*, 612–637.

*Barron, B. (2000). Problem solving in video-based microworlds: Collaborative and individual outcomes of high-achieving sixth-grade students. *Journal of Educational Psychology, 92*(2), 391–398.

Biederman, I., & Shiffrar, M. M. (1987). Sexing day-old chicks: a case study and expert systems analysis of a difficult perceptual-learning task. *J Exp Psychol Learn Mem Cogn, 13*(4), 640–645.

Blumenfeld, P. C., Marx, R. W., Soloway, E., & Krajcik, J. (1996). Learning with peers: from small group cooperation to collaborative communities. *Educational Researcher, 25*(8), 37–40.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. West Sussex: Wiley.

*Brand, S., Reimer, T., & Opwis, K. (2003). Effects of metacognitive thinking and knowledge acquisition in dyads on individual problem solving and transfer performance. *Swiss Journal of Psychology, 62*(4), 251–261.

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: a simple proposal with multiple implications. *Review of Research in Education, 24*, 61–100.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: braIn, mind, experience, and school*. Washington, D.C.: National Academy Press.

*Brodbeck, F.C., & Greitemeyer, T. (2000). Effects of individual versus mixed individual and group experience in rule induction on group member learning and group performance. *Journal of Experimental Social Psychology, 36*(6), 621–648.

Brown, A. L. (1989). Analogical learning and transfer: what develops? In S. Vosniadou & A. Ortony (Eds.), *Similarity and Analogical reasoning* (pp. 369–412). New York: Cambridge University Press.

Bruffee, K. A. (1995). Sharing our toys: cooperative learning versus collaborative learning. *Change: The Magazine of Higher Learning, 27*, 12–18.

Chi, M. T. H., De Leeuw, N., Chiu, M., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognit Sci, 18*, 439–477.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognit Sci, 5*(2), 121–152.

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

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

Cohen, E. G., Lotan, R. A., Scarloss, B. A., & Arellano, A. R. (1999). Complex instruction: equity in cooperative learning classrooms. *Theory into Practice, 38*, 80–86.

Coleman, E. B. (1998). Using explanatory knowledge during collaborative problem solving in science. *Journal of the Learning Sciences, 7*(3–4), 387–427.

*Corliss, S.B. (2005). *The effects of reflective prompts and collaborative learning in hypermedia problem-based learning environments on problem solving and metacognitive skills*. (Unpublished doctoral dissertation). University of Texas at Austin, TX.

*Cuneo, A. (2007). *Examining the effects of collaborative learning on performance in undergraduate mathematics*. (Unpublished doctoral dissertation). Capella University, MN.

Detterman, D. K. (1993). The case for the prosecution: transfer as an epiphenomenon. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: intelligence, cognition, and instruction* (pp. 1–24). Norwood: Ablex.

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

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cogn Psychol, 12*(3), 306–355.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cogn Psychol, 15*(1), 1–38.

Gillies, R. (2004). The effects of cooperative learning on junior high school students during small group learning. *Learning and Instruction, 14*(2), 197–213.

Hedges, L. V., & Olkin, I. (1985). *Statistical model of meta-analysis*. New York: Academic Press.

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychol Methods, 3*(4), 486–504.

Higgins, J., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analysis. *Br Med J, 327*, 557–560. doi:10.1136/bmj.327.7414.557.

*Hudgins, B.B. (1960). Effects of group experience on individual problem solving. *Journal of Educational Psychology, 51*(1), 37-42.

Johnson, D. W., & Johnson, R. T. (1996). Cooperation and the use of technology. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 1017–1044). New York: Macmillan.

Johnson, D. W., & Johnson, R. T. (1999). Making cooperative learning work. *Theory into Practice, 3*(2), 67–73.

Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: social interdependence theory and cooperative learning. *Educational Researcher, 38*(5), 365–379.

Johnson, D. W., Johnson, R. T., & Smith, K. A. (1998). Cooperative learning returns to college: what evidence is there that it works? *Change: The Magazine of Higher Learning, 30*(4), 26–35.

Johnson, D. W., Johnson, R. T., & Stanne, M. B. (2000). Cooperative learning methods: a meta-analysis. Retrieved June 20, 2012, http://www.ccsstl.com/sites/default/files/Cooperative%20Learning%20Research%20.pdf

Johnson, D. W., Maruyama, G., Johnson, R., Nelson, D., & Skon, L. (1981). Effects of cooperative, competitive, and individualistic goal structures on achievement: a meta-analysis. *Psychol Bull, 89*(1), 47–62.

*Kapur, M. (2010). Productive failure in mathematical problem solving. *Instructional Science, 38*(6), 523–550.

*Kirschner, F., Paas, F., & Kirschner, P. (2009). Individual and group-based learning from complex cognitive tasks: Effects on retention and transfer efficiency. *Computers in Human Behavior, 25*(2), 306–314.

*Kirschner, F., Paas, F., Kirschner, P., & Janssen, J. (2011). Differential effects of problem-solving demands on individual and collaborative learning outcomes. *Learning and Instruction, 21*, 587–599.

*Kramarski, B., & Mevarech, Z.R. (2003). Enhancing mathematical reasoning in the classroom: the effects of cooperative learning and metacognitive training. *American Educational Research Journal, 40*(1), 281–310.

*Krause, U-M, & Stark, R. (2010). Reflection in example- and problem-based learning: effects of reflection prompts, feedback and cooperative learning. *Evaluation & Research in Education, 23*(4), 255–272.

*Lambiotte, J.G., Dansereau, D.F., Rocklin, T.R., Fletcher, B., Hythecker, V.I., Larson, C.O., & O'Donnell, A.M. (1987). Cooperative learning and test taking: transfer of skill. *Contemporary Educational Psychology, 12*(1), 52–61.

Lamm, H., & Trommsdorff, G. (1973). Group versus individual performance on tasks requiring ideational proficiency (brainstorming): a review. *European Journal of Social Psychology, 3*(4), 361–388.
Latané, B., Williams, K., & Harkins, S. (1979). Many hands make light the work: the causes and consequences of social loafing. *J Pers Soc Psychol, 37*(6), 822–832.
*Laughlin, P.R., Carey, H.R., & Kerr, N.L. (2008). Group-to-individual problem-solving transfer. *Group Processes & Intergroup Relations, 11*(3), 319–330.
Light, R. J., & Pillemer, D. B. (1984). *Summing up: the science of reviewing research*. Cambridge: Harvard University Press.
Lobato, J. (2006). Alternative perspectives on the transfer of learning: history, issues, and challenges for future research. *Journal of the Learning Sciences, 15*(4), 431–449.
Lou, Y., Abrami, P. C., & D'Apollonia, S. (2001). Small group and individual learning with technology: a meta-analysis. *Review of Educational Research, 71*(3), 449–521.
Manion, V., & Alexander, J. M. (1997). The benefits of peer collaboration on strategy use, metacognitive causal attribution, and recall. *J Exp Child Psychol, 67*, 268–289.
Marton, F. (2006). Sameness and difference in transfer. *Journal of the Learning Sciences, 15*(4), 499–535.
Matthews, R. S., Cooper, J. L., Davidson, N., & Hawkes, P. (1995). Building bridges between cooperative and collaborative learning. *Change: The Magazine of Higher Learning, 27*(4), 34–40.
*McDonald, B.A., Larson, C.O., Dansereau, D.F., & Spurlin, J.E. (1985). Cooperative dyads: Impact on text learning and transfer. *Contemporary Educational Psychology, 10*(4), 369–377.
Michaelsen, L. K., Fink, L. D., & Knight, A. (1997). Designing effective group activities: lessons for classroom teaching and faculty development. In D. DeZure (Ed.), *To Improve the Academy* (Vol. 16, pp. 373–398). Stillwater: New Forums Press and the Professional ad Organizational Development Network in Higher Education.
Mondoux, J., Auderset, P-B, & Dillenbourg, P. (2004). Abstraction and transfer in collaborative learning. Proceedings of the ICLS '04: The 6th international conference on learning sciences (pp. 358–363). International Society of the Learning Sciences.
Moreland, R. L. (2010). Are dyads really groups? *Small Group Research, 41*, 251–267.
*Moreno, R. (2009). Constructing knowledge with an agent-based instructional program: A comparison of cooperative and individual meaning making. *Learning and Instruction, 19*(5), 433–444.
Needham, D. R., & Begg, L. M. (1991). Problem-oriented training promotes spontaneous analogical transfer: memory-oriented training promotes memory for training. *Mem Cognit, 19*, 543–557.
Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. *J Exp Psychol Learn Mem Cogn, 14*, 510–520.
O'Donnell, A. M. (1999). Structuring dyadic interaction through scripted cooperation. In A. M. O'Donnell & A. King (Eds.), *Cognitive perspectives on peer learning* (pp. 179–196). Mahwah: Lawrence Erlbaum Associates.
O'Donnell, A. M. (2006). The role of peers and group learning. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (pp. 781–802). Mahwah: Lawrence Erlbaum Associates Publishers.
O'Donnell, A. M., Dansereau, D. F., Hall, R. H., & Rocklin, T. R. (1987). Cognitive, social/affective, and metacognitive outcomes of scripted cooperative learning. *J Educ Psychol, 79*(4), 431–437.
*O'Donnell, A.M., Dansereau, D.F., Rocklin, T., Lambiotte, J.G., Hythecker, V.I., & Larson, C.O. (1985). Cooperative writing direct effects and transfer. *Written Communication, 2*, 307–315.
*O'Donnell, A.M., Larson, C.O., Dansereau, D.F., & Rocklin, T.R. (1986). Effects of cooperation and editing on instruction writing performance. *Journal of Experimental Education, 54*(4), 207–210.
Okita, S. Y., & Schwartz, D. L. (2013). Learning by teaching human pupils and teachable agents: the importance of recursive feedback. *Journal of the Learning Sciences, 22*(3), 375–412.
*Olivera, F., & Straus, S. (2004). Group-to-individual transfer of learning. *Small Group Research, 35*(4), 440–465.
Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1*(2), 117–175.
Panitz, T. (1999). Collaborative versus cooperative learning: A comparison of the two concepts which will help us understand the underlying nature of interactive learning. Retrieved April 8, 2013, from http://home.capecod.net/~tpanitz/tedsarticles/coopdefinition.htm
Perkins, D. N., & Salomon, G. (1992). Transfer of learning. In T. Husén & T. N. Postlethwait (Eds.), *International encyclopedia of education* (2nd ed., Vol. 11, pp. 6452–6457). Oxford, England: Pergamon Press.
Perkins, D. N., & Salomon, G. (2012). Knowledge to go: a motivational and dispositional view of transfer. *Educational Psychologist, 47*(3), 248–258.
Phelps, E., & Damon, W. (1989). Problem solving with equals: peer collaboration as a context for learning mathematics and spatial concepts. *J Educ Psychol, 81*, 639–646.

Rau, W., & Heyl, B. S. (1990). Humanizing the college classroom: collaborative learning and social organization among students. *Teaching Sociology, 18*(2), 141–155.
*Rebetez, C., Bétrancourt, M., Sangin, M., & Dillenbourg, P. (2010). Learning from animation enabled by collaboration. *Instructional Science, 38*(5), 471–485.
Roscoe, R. D., & Chi, M. T. H. (2008). Tutor learning: the role of explaining and responding to questions. *Instructional Science, 36*(4), 321–350.
Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: rethinking mechanisms of a neglected phenomenon. *Educational Psychologist, 24*(2), 113–142.
Schwartz, D. L. (1995). The emergence of abstract representations in dyad problem solving. *Journal of the Learning Sciences, 4*(3), 321–354.
Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction, 16*(4), 475–522.
Schwartz, D. L., Bransford, J. D., & Sears, D. A. (2005). Efficiency and innovation in transfer. In J. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 1–51). CT: Information Age Publishing.
Schwartz, D. L., Varma, S., & Martin, L. (2008). Dynamic transfer and innovation. In S. Vosniadou (Ed.), *Handbook of conceptual change* (pp. 479–506). Mahwah: Erlbaum.
*Sears, D.A., Packer, E., Hoffa, E., Min, J.W., Sasser, H., & Simison, J. (2011). Collaborating to transfer: are collaboration and transfer synergistic partners? Roundtable presented at the annual meeting of the American Educational Research Association (AERA), New Orleans, LA.
*Sears, D. A., & Pai, H-. H. (2012). Effects of cooperative versus individual study on learning and intrinsic motivation under conditions of reward and reward-removal. *Journal of Experimental Education, 80*, 246–262.
Sears, D. A., & Reagin, J. M. (2013). Individual versus collaborative problem solving: divergent outcomes for accelerated versus traditional students. *Instructional Science, 41*, 1153–1172.
Sharan, Y., & Sharan, S. (1992). *Expanding cooperative learning through group investigation*. New York: Teachers College Press.
Singley, K., & Anderson, J. R. (1989). *The transfer of cognitive skill*. Cambridge: Harvard University Press.
Slavin, R. E. (1977). Classroom reward structure: an analytical and practical review. *Review of Educational Research, 47*, 633–650.
Slavin, R. E. (1983). When does cooperative learning increase student achievement? *Psychol Bull, 94*(3), 429–445.
Slavin, R. E. (1995). *Cooperative learning: theory, research, and practice* (2nd ed.). Boston: Allyn & Bacon.
Slavin, R. E. (1996). Research on cooperative learning and achievement: what we know, what we need to know. *Contemporary Educational Psychology, 21*(1), 43–69.
Slavin, R. E., & Cooper, R. (1999). Improving intergroup relations: lessons learned from cooperative learning programs. *Journal of Social Issues, 55*(4), 647–663.
Souvignier, E., & Kronenberger, J. (2007). Cooperative learning in third graders' jigsaw groups for mathematics and science with and without questioning training. *British Journal of Educational Psychology, 77*(4), 755–771.
Webb, N. M. (1982). Student interaction and learning in small groups. *Review of Educational Research, 52*(3), 421–445.
Williams, K. D. (2010). Dyads can be groups (and often are). *Small Group Research, 41*, 268–274.
*Wirkala, C., & Kuhn, D. (2011). Problem-based learning in K–12 education: is it effective and how does it achieve its effects? *American Educational Research Journal, 48*(5), 1157–1186.
Wood, J. A. (2008). Methodology for dealing with duplicate study effects in a meta-analysis. *Organizational Research Methods, 11*(1), 79–95. doi:10.1177/1094428106296638.