INTRODUCTION

Discussion is... bringing various beliefs together; shaking one against another and tearing down their rigidity. It is conversation of thoughts; it is dialogue—the mother of dialectic in more than the etymological sense.


Discussion is an integral part of our lived experiences. Whether one is sharing a story with a friend, inquiring with the principal about the school’s discipline policy, or debating the reality of the normal distribution with future researchers, individuals are involved in discussion. Discussions in classrooms are fairly open-ended and collaborative episodes of talk among teachers and students, or among students, for the purpose of fostering student thinking, learning, problem-solving, comprehension, or literary appreciation (Wilkinson, 2009). Broadly conceived, classroom discussions can take many forms, including sharing time, content lessons, or even interactions with computers (Cazden & Beck, 2003), and they can involve differing numbers of individuals, such as pairs, small groups, or a whole class, with or without a teacher present (Murphy, Wilkinson, & Soter, 2004).

Different approaches to classroom discussion are structured to serve diverse purposes. Some discussions are focused on a text where the purpose of the talk can vary from the development of critical thinking skills through the sharing of their lived experiences (e.g., Book Club, Raphael & McMahon, 1994) to the thoughtful engagement about a text so as to create a meaningful mental representation (e.g., Questioning the Author, Beck & McKeown, 2006). Other approaches to text-based discussion seek to bridge the sharing of lived-through experiences with a text with the seeking of information, where the purpose of the discussion is to enable students to think critically and analytically about, around, and with a given text (e.g., Quality Talk, Wilkinson, Soter, & Murphy, 2010). Still other discussions do not directly involve text at all, but are focused
on critically analyzing information encountered as part of a class activity (e.g., discussion of observations during or after a science experiment, Mason, 2001; Murphy, 2014). The aforementioned examples all assume mutual interest in the content, allow for give and take, foster probing and exploration, and encourage the collaborative and maternal “giving birth” to newly formed understandings (Dewey, 1916, p. 195).

As Dewey (1916) suggested in Essays in Experimental Logic, discussion also plays a very important role in the development of logical thought. Discussion embodies the very process whereby ideas are brought together and shaken-up and their viability is tested (Dewey, 1916). This testing of ideas encourages doubt and inquiry on the part of discussion participants. Almasi (1996) describes this type of productive discussion as an interactive event in which individuals arrive at new understandings through a collaborative construction of meaning where multiple viewpoints are considered and explored. In such discussions, students have considerable interpretive authority for evaluating the validity and plausibility of individual perspectives as well as any co-constructed understandings (Soter, Wilkinson, Murphy, Rudge, Renninger, & Edwards, 2008).

What is more, Dewey held that participants—as a consequence of participating in this meaning-making experience—internalize the process, and thereafter are able to have such discussions within their own consciousness. As Dewey (1916) states: “The discussion which at first took place by bringing ideas from different persons into contact . . . became a habit of the individual with himself. He became a miniature social assemblage” (pp. 194–195). Similar theoretical perspectives were forwarded by Vygotsky (1978) regarding the development of language and thought.

Specifically, Vygotsky theorized that children developed language as a mechanism for communicating and becoming enculturated into their environment. As children develop physically, cognitively, and socially, the language, tools, and signs of the culture are internalized. Speech that was once external now becomes internal, and language becomes mentally represented as thought. Like Dewey, Vygotsky also contended that along with language, the audience or discourse community is also internalized. Consequently, older learners are able to have internal debates about topics and ideas—debates that were once only possible in the presence of a discourse community. Support for this theoretical position can be found in the research on Collaborative Reasoning (CR; Chinn, Anderson, & Wagger, 2001) where students’ ability to write persuasive arguments independently improves after participation in CR discussions.

The following excerpt is taken from a discussion among a small group of middle-school students and a teacher discussing the moral ambiguities addressed in To Kill a Mockingbird (see Applebee, Langer, Nystrand, & Gamoran, 2003, p. 711). In this excerpt, the teacher asks a pointed question regarding how Bob Ewell is killed.

Teacher: How does Bob Ewell get killed?
Student: Boo Radley [did it].
Teacher: How did you figure out that Boo killed him?
Student: ... But I guess that I thought that the knife ... I really didn’t understand this [part]. I thought it was Boo at the beginning, but then I was not sure.

At this point the teacher intervenes with some scaffolding to help students understand what actually happened in the story. The students turn the focus of the discussion to the issue of the cover-up that follows Ewell’s death (Applebee et al., 2003, pp. 711–712).
Student: It said that he doesn’t want to reveal it to the [sheriff] because . . . it would ruin, you know, Boo’s life.

Student: Right.

Student: Even if he totally [did it].

Student: He’d get all this attention and he couldn’t obviously . . .

Student: No, he wouldn’t be able to [continue to live as before] if they all found out that he did it.

Teacher: Why not?

Student: Well, he’s going to have to go to trial, and uhh . . . all this stuff, and everyone will know about what he has.

Student: I think it’s worth it . . .

Teacher: So you think that Heck Tate was wrong in covering up?

In this excerpted discussion, the many interchanges among students are immediately apparent. Students appear to control the turn taking and the flow of the discussion topic, and manifest the authority to explore varied interpretation of the moral conundrum present in the text. The teacher’s questions are understood by students to be open-ended and authentic. Indeed, following the teacher’s question, the students immediately launch into exploration, probing, and considering possibilities. Also striking in this example is that student talk is more syntactically complete and complex (Applebee et al., 2003). As in a genuine dialogic interchange, the students’ responses build on each another by taking up previous remarks, although not necessarily in the form of a question. The students appear to be genuinely interested in exploring the issue of a cover-up, and the discourse nicely illustrates a situation in which students are exercising interpretive authority. Moreover, the teacher’s response reinforces the authority of the students in interpreting and making meaning of the text (e.g., “So you think that Heck Tate was wrong in covering up?”).

The previous exchange is very different from traditional recitation where the teacher recites what is known about a particular topic or what can be gleaned from a given text (Wilkinson, 2009). In recitation, the teacher holds interpretive authority and controls the talk, the teacher typically talks almost 70% of the time (Cazden, 2001), and exchanges between the teacher and the students follow an IRE (Mehan, 1979) or IRF (Sinclair & Coulthard, 1975) pattern in which the teacher initiates a topic by asking a question; the student responds; and, the teacher evaluates or gives feedback regarding the students response. The IRE or IRF pattern of discourse is often portrayed as counterproductive to student learning or comprehension (e.g., Almasi & Garas-York, 2009; Nystrand, 2006).

In this chapter, rather than portray a given approach as always good or always bad or pit one theoretical perspective against another theoretical perspective, our goal is to take a more pragmatic approach to discussion as an instructional method—an approach in keeping with the writings of early 20th-century American philosophers like William James, John Dewey, and Charles Peirce. In short, our goal is to offer an integrative overview of the theoretical and empirical literature pertaining to classroom discussion with an eye toward the enhancement of student learning. Moreover, given that closely related instructional methods such as cooperative learning (Chapter 18 in this volume), feedback (Chapter 14), or questioning and self-explanation (Chapter 16) are covered in other chapters within this volume, we have chosen to narrow our focus primarily to classroom discussion about, around, or with text. That is, our focus is on text-based discussions, though research and theory on classroom discussions in
general are referenced, as needed, to situate this work. The remainder of this chapter is composed of four major sections, including: (a) a philosophical and sociological history of discussion in which we overview some of the major forces that have helped to shape discussion as an instructional tool; (b) a contemporary theoretical framework where we outline some of the predominant models of classroom discussion about text; (c) current approaches to discussion and key findings in the relevant empirical literature; and (d) emerging trends and future directions in which we consider next steps for teachers and researchers in this important area.

PHILOSOPHICAL AND SOCIOLOGICAL OVERVIEW OF THE HISTORY OF DISCUSSION

The centrality of talk in teaching and learning is as old as language itself. Arguably, what has changed or evolved over time is the emphasis placed on various features of the talk, as well as our understanding of the role that talk plays in student learning. Indeed, there are likely as many perspectives on talk in the annals of history as there are individuals who have written about the topic. Our purpose herein is not to offer a detailed or complete genealogy of the history of discussion as an instructional tool. The exact genealogy of classroom discussion is not overtly clear in the extant literature. This lack of genealogical clarity is likely due to the pervasiveness of talk in every aspect of life, as well as the shared lineages of constructs such as dialogue, dialectic, discourse, rhetoric, and discussion.

Our intention is to highlight philosophers and theorists from history whom we have judged to be key theoretical ancestors of classroom discussion. To us, the theoretical leanings and writings of these ancestors are readily apparent in contemporary theories and models of classroom discussion. Certainly, our selection is not exhaustive nor is it meant to be irrefutable. The forbearers of discussion that we selected include: ancient and medieval philosophers, 16th-century empiricists, and late 19th-century German and American social theorists.

Ancient and Medieval Thinkers: The Birth of Dialectical Reasoning

Explorations of the various types of talk and their role in learning, understanding, and remembering are evident in the texts of ancient Greek philosophers circa 350 BC (e.g., Cornford, 1935). The Socratic method or elenchus, often illustrated in Plato’s dialogues such as the Meno or the Symposium, is perhaps the paragon discussion method. An elenchus discussion is one in which an individual poses an initial or primary question (e.g., What is knowledge?) to another individual who responds to the question, usually with a statement. In most of the Platonic dialogues, Socrates serves as the questioner and some less knowledgeable individual (e.g., Theaetetus) serves as the responder. After this initial exchange, the questioner asks a series of follow-up or secondary questions in order to test the truth value of the responder’s initial statement. In essence, the elenchus is a form of intricate refutation or cross-examination (Robinson, 1953). In many of Plato’s dialogues (e.g., Meno), the primary question aligns very much with what Nystrand (1997) refers to as an authentic question pertaining to an issue of genuine interest and importance, whereas the secondary questions appear much more in line with test questions where there is a predetermined answer.

The overarching purpose of the discourse in these early dialogues is to lay bare the ignorance of the responder regarding the topic—a topic that the responder initially
answered with some assurance. Importantly, however, Socrates thought of himself as midwife to the intellectual progeny of those whom he engaged in dialogue (Woods & Murphy, 2001). In the *Theaetetus*, Socrates states that through his intellectual exchange he could “prove by every test whether the offspring of a young man’s thought [was] a false phantom or instinct with life and truth” (Bostock, 1988, p. 150c). Socrates maintained that only by realizing one’s ignorance could the individual have any hope of gaining a true understanding of the topic.

One of the primary criticisms levied against this type of discussion is that the refutation illuminates that the initial statement of the responder is incorrect but fails to explain *why* the reasoning is flawed (Robinson, 1953). In the later dialogues, Plato alters his method of elenchus to be more integrated into the dialectic, the exchanges were friendlier, and both members of the discussion exhibited increased participation. The questioners more readily acknowledged their attempts to refute the initial claim and the responders attempted to defend their claims (e.g., Gadamer, 1980).

This more open type of discussion laid the groundwork for Aristotle’s syllogistic rules in *Topics* and his notions of proofs in *Rhetoric*. Aristotle defined rhetoric as “the faculty of discovering in the particular case what are the available means of persuasion” (Cooper, 1932, p. 7). Although persuasion is often construed as outside the purview of educational settings and classroom discussions, Aristotle considered persuasion to be a particularly effective means of instruction, particularly when individuals failed to be convinced by scientific arguments (Cooper, 1932). In *Rhetoric*, Aristotle explained that the essence of persuasion lies in the nature of the arguments or proofs forwarded in the text (i.e., oral or written). He described two types of proofs: artistic proofs (i.e., evidence or arguments created by the author) and nonartistic proofs (e.g., laws, witnesses, or contracts). Whereas personal testimony, written contracts, laws, and the like (i.e., nonartistic proofs) are present at the outset of any oratory, artistic proofs must be supplied or invented by the speaker. Artistic proofs must appeal to the knowledge and beliefs of the audience. This can be accomplished in a number of ways. A speaker might gain the confidence of the listener based on personal characteristics as portrayed in speech or text (i.e., ethos) or the speaker may choose to elicit a particular attitude in the listener (i.e., pathos). Still another approach by the speaker would be to present a strong, cogent argument(s) that taps the present knowledge and beliefs of the receiver.

Heavily influenced by Aristotle’s writings, St. Thomas Aquinas was a key figure in the 13th-century Scholastic method of teaching as is evidenced in the *Summa Theologica* (1947). The Scholastic method of teaching emphasized dialectical reasoning in which the primary activity was to find answers to important questions by scouring multiple, usually conflicting, texts and seeking to resolve differences in the texts through philosophical and logical analysis. It is this style of teaching that formed the basis of Aquinas’s philosophy of education, and provided the template for his written arguments. Rather than laying the truth out for the students, Aquinas thought that the process of presenting the question, then introducing arguments for and against, would not only provide a model of academic reasoning for the students but also strengthen the listeners’ understanding of all facets of the issue raised.

Whether it be the elenchus enacted by Socrates in Plato’s later dialogues, Aristotle’s artistic proofs, or the scholasticism of St. Thomas Aquinas, these scholars laid the groundwork for later discussion models that emphasized critical and analytic analyses of talk and text. These ancient scholars encouraged those they taught to grapple with difficult questions and conflicting positions and sources, and talk was one of the
primary mechanisms through which students were encouraged to consider multiple perspectives and the truth value of what was commonly accepted.

17th-Century Empiricists: Sense Experiences Precede Reasoning

The theoretical roots of much of the research on classroom discourse (e.g., Graesser, Gernsbacher, & Goldman, 2003; Kintsch, 2005), particularly cognitively oriented investigations, can be traced to 17th-century British empiricists. Similarly, a number of approaches to text-based discussion (e.g., Instructional Conversations, Goldenberg, 1993) emphasize the need for students to build accurate mental models of the text, and classroom discussion is seen as a mechanism for enhancing students’ ability to build accurate mental representations. As such, we think it is important to look to the philosophical past in order to explore the theoretical roots of these more cognitively oriented, knowledge-focused models.

In the broadest sense, empiricism refers to a philosophical position in which *aposteriori* reasoning (i.e., reasoning from sense experience or observation) provides a mechanism for obtaining substantive knowledge and truths about the world (Locke, 1917). Radical empiricists understand sense experience to be our sole source of knowledge when it comes to the nature of the world (Hume, 1748/1910). It is not that such knowledge is automatic, but rather that deductions or inferences must be based in sensory experience. Empiricism found strong supporters among anti-rationalists of the 17th century, including John Locke, George Berkeley, and David Hume. Among Locke’s (1917) many contributions to empiricism is the notion that children are born without innate ideas and that “perception is the inlet to all knowledge” (p. 72). That is, the mind is a tabula rasa or blank slate on which sensory experiences are imprinted.

Locke also claimed that remembering or reviving memories requires active processing and that the building of complex or compound ideas involves discerning and clarifying among many perceptions. Without the refreshing of both simple and complex ideas, these understandings will decay and be forgotten. Classroom discussion of a text allows students to both revisit the perceptions they acquired during reading, and allows them to gain additional relevant perceptions from the other members of the discussion.

Although contemporary research in educational and cognitive psychology has questioned the notion of pure sensory imprinting, the fact remains that much of what Locke (1917) proposed in his essays on human understanding served to undergird positivistic and post-positivistic research in various realms of psychology. Moreover, as we have illustrated above, some of the basic tenets of Locke’s perspectives on human learning also undergird aspects of contemporary models of classroom discussion about text, particularly those approaches emphasizing the building of mental models of text through discourse (e.g., Beck & McKeown, 2006).
19th-Century Philosophers and Sociologists: The Social Awakening

Although the literatures from the late 19th and early 20th centuries are replete with the writings of scholars who have likely influenced the nature of contemporary models of classroom discussion of text, we have chosen to highlight just a few individuals that we feel brought the importance of the social and cultural condition to the forefront. In turn, the writings of these individuals radically influenced approaches to classroom talk about text. Among those individuals are Karl Marx, noted author of the *Communist Manifesto* (Marx & Engels, 2004), and Charles Cooley, most known for his concept of the looking-glass self (1902). Both in their native countries and abroad, these individuals shed light on the interplay between the individual and society and culture. Arguably, one of the basic tenets underlying text-based discussion models is that students benefit from talk about text because it exposes students to diverse perspectives on a given text. That is, students’ interpretations, explanations, and understandings are colored by the social and cultural experiences that they bring to the reading of the text.

Indeed, the rise of sociocultural theory, which serves as a backbone of many contemporary classroom discussion models in the late 20th century, has been attributed to pragmatist philosophers and anthropologists such as Dewey and Mead, as well as Marxian social theory through the work of Soviet scholars like Vygotsky and Luria (Cole, 1996; Giddens, 1979). Marxian social theory held that the existence of humans is fundamentally based on the ways that humans and nature interact; it is an intricate dance in which the “knower” and the “to be known” mutually adapt. In *Grundrisse*, Marx (1973) states: “Society does not consist of individuals but expresses the sum of interrelations, the relations within which these individuals stand” (p. 265). What matters most in the Marxist paradigm is not society or the individual, but the relations between individuals.

Sociocultural theory has also been traced to Cooley’s writings on human nature and the social order. In fact, Cooley (1902) has been credited as introducing one of the primary tenets of sociocultural theory; that is, the claim of inseparability of the individual and society (Sawyer, 2002). Cooley (1902), in his classic treatise on human nature and the social order, wrote: “Society and individuals do not denote separable phenomena” (pp. 1–2). It is from this notion of the collective that the idea of co-construction of meaning emerges and upon which the role of the community as the repository of legitimate cultural practices takes form (Lave & Wenger, 1991). Indeed, as Marx (1973) suggests, the same object can appear very different and necessarily hold a differential value for individuals from varying social classes.

It is this diversity of thinking that is reflective of one’s unique life experiences that is often manifested during discussions of text. In fact, some researchers of text-based discussion would contend that the understandings of a text that emerge from a group discussion are representative of this inseparability. That is, the meanings cannot be ascribed to any single individual because individual and social thinking are indivisible in a classroom discussion. Indeed, although neither Marx nor Cooley wrote explicitly about classroom discussion, the emphasis that these scholars placed on the social and cultural aspects of human existence is reflected in contemporary models of classroom discussion. This is particularly the case for those approaches to discussion of text that privilege students’ aesthetic and emotive responses to the text and emphasize the importance of breaking down the traditional classroom power structure where the teacher is the gatekeeper to knowledge (e.g., *Book Club*; Raphael & McMahon, 1994).
CONTEMPORARY THEORETICAL FRAMEWORKS

The theory underlying the use of discussions as an instructional tool to improve learning from text derives from cognitive, social constructivist, sociocultural, and dialogic perspectives on learning and teaching. Historically, the cognitive theory preceded the social constructivist and sociocultural theory, so we have chosen to present the theoretical frameworks in that order. Order is in no way meant to imply importance, as our perspective is that there is much to be gleaned from the extant literatures in each of these areas.

Cognitive Theory

Although many contemporary researchers emphasize the sociocognitive and sociocultural theoretical underpinnings of classroom discussion (e.g., Almasi & Garas-York, 2009; Cazden & Beck, 2003), there also exists a rich cognitive theoretical literature pertaining to the ways students process and make meaning from discourse and text (Graesser, Gernsbacher, & Goldman, 2003). Cognitive-based theoretical models, such as the Construction-Integration model (Kintsch, 1988, 1998) or the Concurrent, Capacity-Constrained Activation-Based Production System (3CAPS, Carpenter, Miyake, & Just, 1994; Just & Carpenter, 1992), are being applied to classroom discourse to better understand the mechanisms at play when students attempt to mentally represent and process the various forms of discourse and text present within a classroom environment.

Such cognitive models are also being employed to examine how various mechanisms and processes can be used to predict or explain student learning and change (Graesser, Swamer, Baggett, & Sell, 1996). Indeed, the purpose of these models is to understand the ways in which social interaction promotes individual reasoning and knowledge acquisition (Piaget, 1928) by accounting for both learner characteristics such as prior knowledge, working memory capacity, or beliefs, and features of the discourse or text, including source credibility, coherence, or rhetorical structure (e.g., Kintsch, 1988, 1994; Goldman, Varma, & Coté 1996; Murphy & Alexander, 2004). Although it is beyond the scope of the present chapter to overview all of the various cognitive discourse models, it is important nonetheless to examine some common themes and assumptions across many cognitive models of discourse.

There are a number of commonalities across the more prominent models of discourse and text comprehension (Foltz, 2003). One key commonality is the assumption that discourse comprehension requires that the message be perceived and internalized by the receiver in the form of mental representations (e.g., propositions, semantic networks, or scripts). Discourse researchers have embraced propositional representations more than other representational possibilities (e.g., Kintsch, 1998). In discourse and text models, propositions are the smallest unit assumed to contain meaning and are usually comprised of a predicate (i.e., verb or adjective) and at least one argument (e.g., a noun or another proposition; van Dijk & Kintsch, 1983).

Because propositions are the smallest meaningful unit of discourse, they can be incorporated through processing into other forms of representations, including scripts or schemas. Moreover, although propositional representations capture the meaning of a discourse segment, they do not retain the surface or sentence structure of the message (see Kintsch, 1998). Studies have shown that when cued to a given proposition, individuals are more likely to remember information from within that proposition than information from other propositions from the same sentence.
A second commonality across many cognitive discourse models is that during the encoding of discourse or text, individuals construct increasingly complex levels of cognitive representation (Zwaan & Singer, 2003). At least three levels of increasingly complex representation are widely accepted within the field of cognitive models of discourse comprehension (van Dijk & Kintsch, 1983). Among the levels of discourse processing are the surface form or surface code that refers to a record of the exact wording and syntax of the discourse or text. Generally, the surface form remains in memory for only a few seconds (Gernsbacher, 1985; Jarvella, 1971). Research indicates that retention can be mediated by expectations or pragmatic content, like a joke (e.g., Zwaan, 1996). For example, in hearing or reading a song, when the actual words are of particular importance, listeners generally retain the actual text longer than would be expected for other text genres (Rubin, 1995).

The next level of representation is the textbase, which contains the semantic meaning or explicit propositions from the text. These propositions are generally understood to be networked into a simplified form that retains meaning but loses the surface code syntactic details. As more text is read, the individual adds to the propositional network in working memory (Fletcher & Bloom, 1988). Given the initial processing, the textbase has been shown to be retained in memory for several minutes. Of note is that while van Dijk and Kintsch (1983) separated these first two levels of representation, other researchers have chosen to collapse across these levels (e.g., Johnson-Laird’s propositional representation; 1996).

The final level proposed by van Dijk and Kintsch (1983) is the situation model or the mental model. In creating the situation model, the individual combines the textbase propositions with propositions from long-term memory that serve to provide a fuller and richer understanding of the situation portrayed in the discourse or text (Foltz, 2003). The construction of a sufficient situation model requires an adequate level of prior knowledge or world knowledge pertinent to the message being comprehended. This deeper level of processing also results in longer retention of the situation model in memory. Shallow comprehension requires only the surface code and textbase, whereas deeper comprehension requires the creation of a cohesive, situation model of the text (e.g., Coté, Goldman, & Saul, 1998; Verhoeven & Graesser, 2008; Zwaan et al., 1995). Of importance is the caveat that there remains some debate concerning the extent to which situational models are comprised primarily of propositions or of some combination of propositions and images (Paivio, 2008).

The final commonality across many cognitive models employed to understand text and discourse comprehension pertains to the mechanisms involved in students’ processing of what they hear and read. That is, how exactly are students able to assign meaning to discourse or text? Until more recently, there were at least two competing perspectives on this issue. In the memory-based view of comprehension, it was assumed that the memory-based processes associated with reading took place autonomously and passively, and that activation was the result of the signaling from the discourse or text input and its association to other information, such as prior knowledge (e.g., McKoon, Gerrig, & Greene, 1996).

In stark contrast, the constructionist perspective assumed that learners must actively engage in the search for meaning while participating in discourse or reading a text (Graesser, Singer, & Trabasso, 1994). During the activity of reading, students were assumed to be goal directed, monitoring coherence of the various levels of representation and seeking explanations for outcomes and inferences within the text (Zwaan & Singer, 2003). Interestingly, there exists convincing empirical evidence in support of
both positions (see van den Broek, Rapp, & Kendou, 2005). Over the last decade, however, the tensions between these two perspectives have given way to cognitive models that integrate the memory-based position with the constructionist position in which some processes take place passively while others require active, conscious processing of the text (Verhoeven & Graesser, 2008). The Landscape Model proposed by van den Broek and colleagues (2005) serves as an example of an empirically supported integrated model. What is important about these cognitive theories of text and discourse comprehension is that they provide a theoretical mapping for researchers to create discussion approaches that will enhance these comprehension processes. As Locke suggested so many years ago, such models would enhance the construction process and the building of meaningful mental models that can aid in comprehension and retention of content.

Social Constructivist and Sociocultural Theory

Social constructivist and sociocultural theory are routinely invoked to explain the role of discussion in promoting students’ understandings of discourse and text (Murphy et al., 2009). As perhaps the eminent social constructivist, Vygotsky’s theoretical and empirical writings have had a profound effect on models of classroom discussion. Vygotsky (1934/1986) conceived of learning as a culturally embedded and socially mediated process in which discourse plays a primary role in the creation and acquisition of shared meaning making. Moreover, children develop the skills and abilities needed to read through the participation in literacy-rich environments. Within these environments, children are apprenticed into the literate community through authentic, real world participation with more knowledgeable others, including parents, teachers, or more capable peers. Importantly, while the skill development and apprenticeship took place on the social plane, Vygotsky (1978) also conceptualized reading and writing as higher-order psychological processes. Essentially, Vygotsky understood these higher order psychological processes as emerging out of students’ participation in literacy communities (Tharp & Gallimore, 1988).

Vygotsky argues that a child’s development cannot be understood by a study of the individual. We must also examine the external social world in which that individual life has developed . . . Through participation in activities that require cognitive and communicative functions, children are drawn into the use of these functions in ways that nurture and “scaffold” them. (Tharp & Gallimore, 1988, pp. 6–7)

As was advanced in Cooley’s (1902) early writings on nature and the social order, students involved in talk about text necessarily bring to the discussion unique cultural and social values, differential levels of prior knowledge, varied background experiences, and potentially disparate assumptions about learning and text. Through interactions in the discussion, learners are exposed to, evaluate, and possibly incorporate ways of thinking about the text that are very different from their individually contrived understandings. When students interact with others in a group in deep and meaningful ways, the outcomes or results that are produced are beyond the abilities and dispositions of the individual students who comprise the group (Wertsch, Del Rio, & Alvarez, 1995). As Dewey (1916) suggested, these new abilities and dispositions are then internalized by learners and can be transferred to other activities that involve independent learning or problem solving (Anderson et al., 2001; Hatano, 1993).
Another key assumption of the social constructivist and sociocultural theory is that reasoning is inherently dialogical and that growth and development are encouraged through dialogic reasoning (Bahktin 1981; Vygotsky, 1978). That is, one’s reasoning is necessarily a response to what has been said or experienced as well as an anticipation of what will be said in response. It is not so much that one cannot reason individually but rather that reasoning is mediated by prior experiences and the anticipation of future social experiences. In short, the reasoning process, like the social environment, is dynamic and relational. In the context of text comprehension and classroom discussion, the dialogic process is negotiated and sustained through interpretations of text, high-level reasoning, and standards of interaction that govern group behavior. It is for this reason that approaches like Collaborative Reasoning are structured so as to encourage learners to consider their own perspective, as well as the perspectives of their peers (Anderson et al., 2001; Reznitskaya et al., 2001). As Vygotsky suggests, these types of social interactions encourage cognitive growth. “[Growth is] more likely when one is required to explain, elaborate, or defend one’s position to others, as well as to oneself; striving for an explanation often make a learner integrate and elaborate in new ways” (Vygotsky, 1978, p. 158).

Our goal in this section has been to overview the various contemporary theories influencing models of classroom discussion designed to enhance text comprehension. We contend that deeply understanding the ways in which classroom discourse enhances student learning requires an integration of the cognitive, social constructivist, and sociocultural theoretical perspectives. Simply put, it is not enough to embrace the outward language, tools, and signs that students make use of during discussion without considering how these exchanges are internalized, encoded, and processed by the learner, or the manner in which students justify their views or substantiate their claims. To date, however, the majority of approaches to classroom discussion attend more to the socially shared and socially constructed nature of learning than to the cognitive internalization of what is being shared. In the section that follows, we overview trends and issues at play in contemporary research on classroom discussion.

CURRENT APPROACHES TO DISCUSSION

An issue confronting the field is what the different approaches to conducting discussions about text have to offer teachers and their students. As indicated earlier, there are a multitude of approaches to conducting text-based discussions, and it is difficult for educators to make informed decisions about when, why, and how to use them. A related issue confronting the field is the strengths and weaknesses of the approaches in terms of their impacts on students’ learning and comprehension of text. In this section, we identify the major approaches to conducting discussions about text and attempt to make sense of them by describing their similarities and differences. We then examine what is known about their effects on students’ learning and comprehension.

Making Sense of the Approaches

Wilkinson, Soter, and Murphy (2007) conducted a synthesis of research on classroom discussions about text and identified nine major approaches. A meta-analysis of the effects of discussion on students’ high-level comprehension of text (Murphy, Wilkinson, Soter, Hennessey, & Alexander, 2009), as well as an extensive analysis of discourse
Instruction Based on Discussion

from students’ classroom discussions of text (Soter et al., 2008), were conducted as part of the larger project. To qualify as a major discussion approach in the meta-analysis and discourse analysis, Wilkinson and colleagues (2007) stipulated that the approach had to demonstrate consistency of application—it had to look the same wherever it was implemented—and have an established place in educational research or practice based on a record of peer-reviewed, empirical research published since 1970. Using these criteria, they identified nine major approaches: Instructional Conversations (Goldenberg, 1993), Junior Great Books Shared Inquiry (Great Books Foundation, 1987), Questioning the Author (Beck & McKeown, 2006), Collaborative Reasoning (Anderson, Chinn, Waggoner, & Nguyen, 1998), Paideia Seminars (Billings & Fitzgerald, 2002), Philosophy for Children (Sharp, 1995), Book Club (Raphael & McMahon, 1994), Grand Conversations (Eeds & Wells, 1989), and Literature Circles (Short & Pierce, 1990).

There are other approaches to conducting classroom discussions about text. These include Conversational Discussion Groups (O’Flahavan, 1989), Dialogical-Reading Thinking Lessons (Commeeyras, 1993), Idea Circles (Guthrie & McCann, 1996), and Point-Counterpoint (Rogers, 1990). These approaches have some prominence in the field but they did not meet Wilkinson et al.’s (2007) criteria because there is relatively little empirical research on them. There are also various instantiations of literature discussion groups based on reader-response theory (see Gambrell & Almasi, 1996), discussion-based envisionments of literature (Langer, 1993; 1995, 2001), and instructional integrations of writing, reading, and talk (Nystrand, Gamoran, & Carbonaro, 2001; Sperling & Woodlief, 1997). These approaches have also received attention in the research literature but they do not show the consistency of application necessary to consider them as distinct approaches (Wilkinson et al., 2007). There is also an approach called Accountable Talk, developed by Lauren Resnick and colleagues (Michaels, O’Connor, & Resnick, 2008; Michaels, O’Connor, Hall, & Resnick, 2002). Accountable Talk comprises a set of conditions for productive conversation in academic contexts and forms part of the New Standards Project (Resnick, 1999; Resnick & Hall, 1998). Wilkinson et al. did not include it in their review because it is not specifically designed for conducting text-based discussions (although it has applicability as an approach for promoting reading comprehension; see Wolf, Crosson, & Resnick, 2005).

The similarities and differences in the nine major approaches identified by Wilkinson et al. (2007) can best be described by characterizing them on various dimensions. Table 20.1 describes the nine approaches in terms of several aspects of a discussion: the stance toward the text, who chooses the topic for discussion, who has interpretive authority, who controls turn taking, who chooses the text, the size of the discussion group, and whether the group is peer- or teacher-led. Stance toward the text is largely established by the teacher and classroom context and is usually categorized in terms of an aesthetic, efferent, or critical-analytic stance. An aesthetic or, more appropriately, expressive (Soter, Wilkinson, Connors, Murphy, & Shen, 2010), stance privileges a reader’s affective response to the text, her spontaneous, emotive connection to all aspects of the textual experience (Rosenblatt, 1978). An efferent stance privileges a more utilitarian response to the text—reading for the purpose of acquiring and retrieving information. The focus is on “the ideas, information, directions, conclusions to be retained, used, or acted on after the reading event” (Rosenblatt, 1978, p. 27). A critical-analytic stance privileges a more objective, critical response in which the reader interrogates or queries the text in search of the underlying arguments, assumptions, worldviews, or beliefs (Wade, Thompson, & Watkins, 1994).
As can be seen from Table 20.1, two important dimensions on which discussions vary are the dominant stance toward the text and the degree of control of the discussion exerted by the teacher versus the students (cf. Chinn et al., 2001). Wilkinson et al. (2007) noticed that these two dimensions of discussions are related. Discussions that give prominence to an aesthetic stance toward the text tend to be those in which students have the greatest control over topic, interpretive authority, turn taking, and choice of text. These approaches include Literature Circles, Grand Conversations, and Book Club. Conversely, discussions that give prominence to an efferent stance tend to be those in which teachers have the greatest control. These approaches include Questioning the Author, Instructional Conversations, and Junior Great Books Shared Inquiry. Discussions that give prominence to a critical-analytic stance tend to be those in which teachers and students share control. In these approaches, the teacher has considerable control over the choice of text and topic, but students have considerable interpretive authority and control of turns. These approaches include Paideia Seminars, Collaborative Reasoning, and Philosophy for Children.

Another dimension on which discussions vary is the size of the discussion group (small-group versus whole-class). As shown in Table 20.1, Questioning the Author, Junior Great Books Shared Inquiry, Paideia Seminar, and Philosophy for Children tend to use whole-class arrangements or, at least, discussions with large groups of about 10–14 students. The other approaches tend to use smaller groups. The available evidence suggests that smaller groups are better for discussion, although they should not be so small as to limit the diversity of ideas needed for productive discussion (Wiencek & O’Flahavan, 1994). Morrow and Smith (1990), in a study of kindergarten students who engaged in discussions of stories that were read aloud, reported benefits of small-group discussions compared to one-on-one discussions with the teacher or whole-class discussions. Sweigart (1991), in a study of 58 twelfth-grade students, found that student-led, small-group discussions produced greater effects on students’ recall and understanding of essays they had read than did lecture or whole-class discussions. Smaller groups provide more opportunities for students to speak, interact, and exchange points of view in discussion, thus contributing to greater knowledge and understanding of the text and of how to make sense of text.

Yet another dimension on which discussions vary is whether they are peer- or teacher-led. Book Club discussions (those that occur in small groups rather than the whole class “Community Share”) and many Literature Circles are peer-led, whereas the other types of discussions are teacher-led. The relative merits of the two formats have been the subject of some debate and available research has not yielded a definitive answer on the issue. Peer-led discussions can enable students collectively to explore topics more fully and to have more control and interpretive authority (Almasi, 1994). However, in teacher-led discussions, the teacher can play an important role in keeping students on topic, fostering norms for productive discourse, and modeling and scaffolding talk to enhance the quality of learning opportunities for students (O’Flahavan, Stein, Wiencek, & Marks, 1992; Michaels, O’Connor, & Resnick, 2008; Soter et al., 2008; Wells, 1989). Most likely the key question is not so much whether the teacher or students should lead the group as it is how much structure and focus is provided while giving students the flexibility and responsibility for thinking and reasoning together (cf. Mercer, 1995). Productive discussions need to be structured and focused but flexible enough to foster generative learning—and these can peer- or teacher-led.

Rooted in the findings of their extensive synthesis of the literature (Wilkinson et al., 2007), meta-analysis of effects of text-based discussion on high-level comprehension
<table>
<thead>
<tr>
<th>Aspect</th>
<th>LC</th>
<th>GC</th>
<th>BC</th>
<th>QtA</th>
<th>IC</th>
<th>JGB</th>
<th>PS</th>
<th>CR</th>
<th>P4C</th>
<th>QT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expressive Stance</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Low to Medium</td>
<td>Low</td>
<td>Low</td>
<td>Medium</td>
<td>Medium</td>
<td>High</td>
<td>Medium to High</td>
</tr>
<tr>
<td>Efferent Stance</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium to High</td>
</tr>
<tr>
<td>Critical-Analytic Stance</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Control of Topic</td>
<td>Students</td>
<td>Students</td>
<td>Students</td>
<td>Teacher</td>
<td>Teacher</td>
<td>Teacher</td>
<td>Teacher</td>
<td>Teacher</td>
<td>Teacher</td>
<td>Teacher</td>
</tr>
<tr>
<td>Interpretive Authority</td>
<td>Students &amp; Teacher</td>
<td>Students &amp; Teacher</td>
<td>Students</td>
<td>Teacher</td>
<td>Teacher</td>
<td>Teacher</td>
<td>Students &amp; Teacher</td>
<td>Students</td>
<td>Student</td>
<td></td>
</tr>
<tr>
<td>Control of Turns</td>
<td>Students</td>
<td>Students</td>
<td>Students</td>
<td>Teacher</td>
<td>Students &amp; Teacher</td>
<td>Teacher</td>
<td>Students &amp; Teacher</td>
<td>Students</td>
<td>Student</td>
<td></td>
</tr>
<tr>
<td>Chooses Text</td>
<td>Students</td>
<td>Teacher</td>
<td>Students</td>
<td>Teacher</td>
<td>Teacher</td>
<td>Teacher</td>
<td>Teacher</td>
<td>Teacher</td>
<td>Teacher</td>
<td>Teacher</td>
</tr>
<tr>
<td>Small group or Whole class</td>
<td>Small group</td>
<td>Small group</td>
<td>Small group</td>
<td>Whole class</td>
<td>Small group / Whole class</td>
<td>Whole class</td>
<td>Small group</td>
<td>Whole class</td>
<td>Small group</td>
<td>Whole class</td>
</tr>
<tr>
<td>Peer or Teacher Led</td>
<td>Peer</td>
<td>Teacher</td>
<td>Peer</td>
<td>Teacher</td>
<td>Teacher</td>
<td>Teacher</td>
<td>Teacher</td>
<td>Teacher</td>
<td>Teacher then Peer</td>
<td>Teacher</td>
</tr>
</tbody>
</table>

Note: LC = Literature Circles, GC = Grand Conversations, BC = Book Club, QtA = Questioning the Author, IC = Instructional Conversations, JGB = Junior Great Books, PS = Paideia Seminar, CR = Collaborative Reasoning, P4C = Philosophy for Children, QT = Quality Talk.
(Murphy et al., 2009), and analysis of talk in text-based discussions (Soter et al., 2008), Wilkinson and colleagues developed an additional model of discussion that they aptly named Quality Talk—a moniker that highlights the importance of the nature and quality of the talk in classrooms (Wilkinson, Murphy, & Soter, 2010). Quality Talk was created by combining the most effective aspects of the aforementioned discussion approaches with the specific intention of enhancing students’ ability to talk about, around, and with text (i.e., high-level comprehension) and to engage text in critical and analytic ways. The features of Quality Talk are overviewed in Table 20.1. Revisions and improvements of the Quality Talk model also employ epistemic discourse engagement as a mechanism to improve students’ argumentation and reasoning (Murphy, Greene, & Firetto, 2015).

EXAMINING EFFECTS OF DISCUSSION ON LEARNING

Research has examined whether discussion improves the learning and comprehension of the texts that were the subject of discussion as well as the learning and comprehension of new texts not discussed in class or performance on new but related tasks. It stands to reason that enabling students to engage in discussions about texts should improve their comprehension of those texts. For example, in an experimental study, Fall, Webb, and Chudowsky (2000) analyzed tenth-grade students’ performance on language arts tests in which students either discussed or did not discuss a story they were required to read and interpret. Their results showed that allowing students to engage in a 10-minute discussion of the story in three-person groups improved students’ factual knowledge of the story as well as their understanding of characters’ motives and feelings and the story’s theme. Similarly, in a quasi-experimental study, Van den Branden (2000) showed that allowing fifth-grade Dutch students, many of whom were non-native speakers, to negotiate the meaning of story by means of discussion improved their comprehension of the main ideas relative to that of students who simply read the story.

The more interesting and important issue for educational research is whether discussion enables students to acquire the habits of mind to transfer their comprehension and learning capabilities to new texts and novel tasks. For example, fourth- and fifth-grade students who participated in Quality Talk evidenced statistically significant growth in comprehension, irrespective of whether the comprehension assessment was conducted on text the students had discussed or not (Murphy et al., 2015) as well as on transfer measures of persuasive writing (Firetto, Murphy, Greene, Li, Wei, Croninger, & Montalbano, 2015). In this section, we examine the role of discussion in enhancing students’ comprehension and learning outcomes, giving particular attention to the results of studies in which researchers have assessed the effects of discussion on measures that are independent of the texts that were discussed. We first consider correlational studies, then quasi-experimental and experimental studies in general, then the subset of quasi-experimental and experimental studies that addressed the issue of transfer.

Correlational Studies

Applebee et al. (2003) conducted a large-scale correlational study of the relationship between discussion and students’ literary performance. They observed instructional practices in 64 middle- and high-school English classrooms across the US on four occasions over the school year. They also assessed the students’ literary performance in
terms of their ability to write essays, scored for levels of abstraction and elaboration, collecting data on 974 students. One of the writing tasks related to a novel, short story, or play that they had studied during the year, but the other task related to writing about a general experience. Their results showed that discussion-based practices, used in the context of academically challenging tasks, were positively related to students’ literary performance on the writing tasks.

Similar results have been reported in other correlational studies. Langer (2001) studied the instructional practices associated with student achievement in 25 middle and high schools, involving 44 teachers and 88 classes. The study involved a nested, multiple-case design comparing practices in schools with higher-than-expected achievement in literacy with those in more typically performing schools. She found that whole-class and small-group discussion was one of the characteristics of instruction in the schools that showed higher-than-expected achievement in reading, writing, and English.

In another correlational study, Taylor, Pearson, Clark, and Walpole (2000) observed the instructional practices in first- through third-grade classrooms in 14 high-poverty schools in Virginia, Minnesota, Colorado, and California and compared the practices in schools categorized as most, moderately, or least effective in promoting student reading achievement. They showed that asking higher-level, aesthetic-response questions in discussions about text was a feature of instruction of the most accomplished teachers and of teachers in the most effective schools (see also Taylor, Pearson, Clark, & Walpole, 1999). In subsequent school change work to promote the cognitive engagement of students in grades 1–5 in high-poverty schools, Taylor, Pearson, Peterson, and Rodriguez (2002, 2003, 2005) again found that the incidence of teachers’ higher-level questions about text predicted students’ end-of-year achievement on a number of measures of reading and writing (though not always on measures of reading comprehension).

**Quasi-Experimental and Experimental Studies**

As mentioned, Murphy et al. (2009) conducted a meta-analysis of 42 quantitative studies of the effects of the nine major approaches to text-based discussions described earlier. They included single-group, pretest-posttest design studies and multiple-group studies and examined effects on measures of teacher and student talk as well as on measures of individual student comprehension and learning outcomes. Results showed that the approaches were differentially effective in promoting comprehension. Many of the approaches were effective at promoting students’ literal and inferential comprehension, especially those that had a more efferent stance toward the text, namely Questioning the Author, Instructional Conversations, and Junior Great Books Shared Inquiry. Some of the approaches were particularly effective at promoting students’ critical-thinking, reasoning, and argumentation about text, namely Collaborative Reasoning and Junior Great Books Shared Inquiry. Another finding from Murphy et al.’s meta-analysis was that increases in student talk did not necessarily result in concomitant increases in student comprehension. Rather, a particular kind of talk was necessary to promote comprehension. This type of talk is characterized by authentic, student-generated questioning, elaborated explanations, and exploratory talk, where students’ hold interpretive authority and the teachers’ use select discourse moves to facilitate the discussion. This is consistent with observations from other research that the success of discussion hinges not on increasing the amount of student talk per se,
but in enhancing the quality of the talk (Wells, 1989). Results of the meta-analysis also suggested that the approaches exhibited greater effects for students of below-average ability than for students of average or above-average ability. Murphy et al. interpreted this finding to mean that the higher-ability students might be able to read a text and think independently about the nuances of meaning even without participating in discussion.

In the discussion that follows, we expand upon the aforementioned general trends emerging from the meta-analytic study conducted by Murphy et al. (2009). In doing so, we offer specific effect sizes emerging from noteworthy studies within each of the given approaches as estimates of the effectiveness of the various approaches to classroom discussion. In Table 20.2, we summarize across the various studies conducted by researchers pertaining to each approach both single-group (i.e., within-subjects designs) and multiple groups (i.e., between-subjects designs). In doing so, we hoped to provide a somewhat comprehensive picture of the nature of the effectiveness of the various prominent approaches to classroom discussion.

**Quasi-Experimental and Experimental Studies Addressing Transfer**

In the balance of this section, we focus on the multiple-group studies that examined the effects of discussion on measures of transfer to new texts or novel tasks. Anderson and colleagues (Dong, Anderson, Kim, & Li, 2008; Kim, 2001; Reznitskaya, Anderson, & Ku, 2007; Reznitskaya et al., 2001) conducted a number of quasi-experimental studies of Collaborative Reasoning with fourth and fifth-grade students. They compared the performance of students who participated in anywhere from 4 to 10 Collaborative Reasoning discussions with that of students in control conditions who received regular classroom reading instruction. Students’ performance was assessed on a persuasive essay-writing task measured by counting the numbers of arguments, counter-arguments, and rebuttals in the essays. Anderson and colleagues’ results showed that the essays of students who participated in the Collaborative Reasoning discussions contained a greater number of argument components than the essays of students in the control conditions. The magnitude of the effects on the total number of argument components was moderate to large with effect sizes ranging from $d = 0.45$ to 0.68 (Reznitskaya et al., 2008). One interpretation of these results is that the students internalized an argument schema from the oral group discussions and transferred this capability to written argumentation performed individually and independently.

The most stringent tests of the benefits of discussions come from quasi-experimental and experimental studies that examined the effects of discussion, relative to a control condition, on norm- or criterion-referenced standardized measures (rather than researcher-developed measures). Bird (1984) conducted a quasi-experimental study of the effects of the Junior Great Books program on fifth-graders’ critical reading and thinking skills. One hundred and eight higher-level readers in four school districts of Monmouth County, New Jersey, participated in the study. One group received instruction full-time in the Junior Great Books program, one group received instruction in a basal reading program, and one group received instruction in a combination of Junior Great Books and a basal program. Students from different districts were given different treatments, and the duration of treatment varied between groups (three to six months with a mean = 4.1 months). At the end of the treatment, students’ performance was assessed on the Worden Test of Critical Thinking and Reading and the Ross Tests of Higher Cognitive Processes. Results showed that students’ full or part-time
participation in Junior Great Books resulted in significant gains in scores on both measures as compared to participation in the basal program only (median $d = 0.46$). There were no significant differences between the scores of the full-time and combination groups.

Heinl (1988) conducted another quasi-experimental study of the effects of the Junior Great books program on fifth graders’ comprehension over a period of six months. Thirty students were allocated to three groups of equal reading ability. One group met

### Table 20.2 Effect Sizes by Construct Comparing Single- and Multiple-Group Studies within Approach

<table>
<thead>
<tr>
<th>Stance</th>
<th>Approach</th>
<th>Grouping</th>
<th>TT</th>
<th>ST</th>
<th>Comp</th>
<th>TE</th>
<th>TI</th>
<th>SI</th>
<th>CT/R</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Multiple</td>
<td></td>
<td></td>
<td>.262</td>
<td></td>
<td></td>
<td>.668</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P4C</td>
<td>Single</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multiple</td>
<td>-.291</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.236</td>
</tr>
<tr>
<td></td>
<td>PS</td>
<td>Single</td>
<td>-.030</td>
<td>-.006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multiple</td>
<td>-.655</td>
<td>.446</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.428</td>
</tr>
<tr>
<td>Efferent</td>
<td>QtA</td>
<td>Single</td>
<td>.098</td>
<td>.330</td>
<td>-.205</td>
<td>.949</td>
<td></td>
<td></td>
<td>2.499</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multiple</td>
<td></td>
<td></td>
<td></td>
<td>.800</td>
<td></td>
<td></td>
<td>.627</td>
</tr>
<tr>
<td></td>
<td>IC</td>
<td>Single</td>
<td></td>
<td></td>
<td>2.735</td>
<td>2.798</td>
<td>2.988</td>
<td></td>
<td>1.263</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multiple</td>
<td>-.408</td>
<td>1.653</td>
<td></td>
<td>.509</td>
<td>.568</td>
<td></td>
<td>.610</td>
</tr>
<tr>
<td></td>
<td>JGB</td>
<td>Single</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.392</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multiple</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.408</td>
</tr>
<tr>
<td>Expressive</td>
<td>LC</td>
<td>Single</td>
<td>-.439</td>
<td>1.637</td>
<td>.633</td>
<td></td>
<td>2.136</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multiple</td>
<td></td>
<td></td>
<td></td>
<td>.114</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GC</td>
<td>Single</td>
<td>.043</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.822</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multiple</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BC</td>
<td>Single</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multiple</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: CR=Collaborative Reasoning, P4C=Philosophy for Children, PS=Paideia Seminar, QtA=Questioning the Author, IC=Instructional Conversation, JGB=Junior Great Books, LC=Literature Circles, GC=Grand Conversation, BC=Book Club, TT=Teacher Talk, ST=Student Talk, Comp=General Comprehension, TE=Text-Explicit Comprehension, TI=Text-Implicit Comprehension, SI=Scriptally-Implicit Comprehension, and CT/R=Critical-Thinking/Reasoning.
every two weeks to discuss stories from Junior Great Books; one group read and wrote summaries of the same stories; and one served as a control group that read materials from the fifth-grade basal series. At the end of the program, students’ performance was assessed on the Iowa Tests of Basic Skills (ITBS). Results showed that there were no significant differences among the three groups on the ITBS reading comprehension post-test. However, there was a significant difference for the low-ability students. The low-ability students who discussed stories from the Junior Great Books program scored higher than the low-ability students in the other two groups on the ITBS reading comprehension post-test. Heinl concluded that the Junior Great Books benefited low-ability students in reading comprehension but did not benefit the high-ability students. It should be noted that all three groups read the materials from the fifth-grade basal series, so the Junior Great Books program was a supplement to students’ regular reading instruction.

Lipman (1975) carried out an early pilot study of Philosophy for Children in a New Jersey school during 1970–1971. Forty middle- and low-income African American students participated. These students were organized into small groups and the groups were randomly assigned to experimental and control conditions. Lipman and two graduate students taught the students in the experimental group according to the Philosophy for Children framework while the control group was given social science instruction. The classes met twice a week for nine weeks. The results of the initial post-test apparently showed that the experimental group students made significant gains over the control group, but the original data were lost, so these results cannot be verified. Two and one-half years later, Lipman administered the reading subtest of the Iowa Test of Basic Skills to the students when they were in seventh grade. Results showed a statistically significant and large difference ($d = 0.56$) between the experimental and control group students in favor of Philosophy for Children. This is the strongest effect found for Philosophy for Children on a standardized test, and the result needs to be interpreted cautiously. It is not known what instruction students received in the two and one-half years intervening between the implementation of the program and the follow-up testing and whether it differed for students in the experimental and control conditions.

Yeazell (1982) conducted another quasi-experimental study of Philosophy for Children. One hundred fifth-grade students from five classes and three schools participated. Four classes were assigned to the experimental group and one class to the control group. The teacher of the control class also taught one of the experimental classes. Students in the experimental condition participated in the Harry Stottlemeier’s Discovery component of the Philosophy for Children program once a week over the school year as a supplement to the regular reading program. Students completed the Comprehensive Test of Basic Skills (CTBS) the year before and at the end of the school year. Students in the experimental group made statistically significant gains on the CTBS over the year whereas students in the control group did not. Those students taught by the same teacher showed the same results; the experimental group made a significant gain in test scores while the control group did not. Yeazell concluded that implementing philosophical discussions into reading curriculum enhanced the reading comprehension of students.

The effect of Philosophy for Children on students’ reading comprehension was also examined in a quasi-experimental study by Banks (1987). The teachers of three grade 2 classes, four grade 4 classes, and two grade 5 classes volunteered to use the Philosophy for Children program over a school year, while equal numbers of teachers at each
grade level, selected by school district administrators, served as the control classrooms. Students completed the California Achievement Test before and after the school year and gain scores were calculated. In reading, results showed no significant differences in gain scores between groups at second grade ($d = -0.18$), a significant difference in favor of Philosophy for Children at fourth grade ($d = 0.46$), and no significant difference at fifth grade ($d = 0.27$). We note that the teachers self-selected into the Philosophy for Children program, and it was reported that implementation of the program varied among teachers.

Firmer conclusions about the effects of Philosophy for Children can probably be drawn from an experimental study conducted by Chamberlain (1993). Eighty fourth- and fifth-grade gifted students in two elementary schools participated in the study with students being randomly assigned to experimental and control group within each class. Students in the experimental condition studied the Harry Stottlemeier’s Discovery component of the Philosophy for Children program five times a week, for one hour each session, for 12 weeks. The students in control groups studied other literature selections for the same amount of time. Students completed the New Jersey Test of Reasoning Skills and the Ross Test of Higher Cognitive Processes at the end of the 12 weeks. Results showed statistically significant differences between groups in favor of Philosophy for Children in scores on the New Jersey Test of Reasoning Skills ($d = 0.29$) but not on the Ross Test of Higher Cognitive Processes ($d = 0.29$). The latter non-significant finding might have been due to ceiling effects on the Ross Test.

Reznitskaya and colleagues (2012) conducted a quasi-experimental study in which they assigned 12 fifth-grade classrooms to either P4C (i.e., Philosophy for Children discussions) or a regular instruction control. Discourse pattern data was collected and analyzed in addition to individual students’ transfer to measures of persuasive writing, oral argumentation, and text recall. Coded discourse of P4C discussions supported the notion that the discussions were inherently dialogic, while the coded discourse in the regular instruction classroom showed that the discussions followed a traditional IRE (i.e., initiate, respond, and evaluate) pattern. Despite these differences in the discourse, there were no differences between students in the two discussion groups on any of the transfer measures. Reznitskaya and colleagues offer several explanations for these unexpected findings, most notably being that students may need more support to transfer from discourse to individual outcome measures.

**SUMMARY**

The findings from the correlational, quasi-experimental and experimental studies reviewed here indicate that there is, at least, a moderate level of evidence that classroom discussions have positive effects on reading and literacy-related outcomes including those that are independent of the texts discussed under certain conditions. The correlational studies are large in scale but suffer from the problems of interpretation typical of such studies, namely that other factors might have been responsible for the observed ‘effects’ or that the causal direction of the effects is the opposite of that assumed (i.e., students who are better readers engage in more discussion-based practices than do other students). The quasi-experimental studies of Collaborative Reasoning by Anderson and colleagues (Dong, Anderson, Kim, & Li, 2008; Kim, 2001; Reznitskaya, Anderson, & Ku, 2007; Reznitskaya et al., 2001) are rigorous in design and show moderate to large positive effects on a researcher-developed measure. The quasi-experimental studies of Junior Great Books by Bird (1984) and Heinl (1988) and
those of Philosophy for Children by Lipman (1975), Yeazell (1982), and Banks (1987) are less rigorous and show small to moderate positive effects on norm- and criterion-referenced measures. Both sets of quasi-experimental studies are open to the interpretation that the observed effects are due to factors specific to the classes or groups under study rather than to the discussion per se. Chamberlain’s (1993) experimental study of Philosophy for Children is probably the most rigorous in design and shows small effects on standardized measures of critical thinking and reading. It might be argued that teacher effects are at work in the quasi-experimental and experimental studies, but this argument becomes less tenable as the number of studies showing positive effects of discussion grows. Taken together, the studies reviewed in this section suggest that classroom discussions about text can enable students to acquire the habits of mind to transfer their comprehension capabilities to new texts and novel tasks.

EMERGING TRENDS AND FUTURE DIRECTIONS

Although discussion is an integral part of our lived experiences, discussions in classrooms seem to be relatively rare. Commeyras and Degroff (1998) surveyed the pedagogical practices of a random sample of 1,519 K–12 literacy teachers and related professionals in the US and reported that only 33% of respondents said they frequently or very frequently had students discuss literature in their classrooms. Similarly, Nystrand (1997) observed the instructional practices of teachers in 58 eighth-grade and 54 ninth-grade language arts and English classes in eight Midwestern communities in the United States. He found that open-ended, whole-class discussion averaged only 52 seconds per class in eighth grade and only 14 seconds per class in ninth grade. By contrast, anecdotal reports suggest that recitation, where teacher and student exchanges follow a traditional IRE pattern, is still pervasive in elementary and high school classrooms (Almasi, 1994; Cazden, 2001; Goldenberg, 1992; Tharp & Gallimore, 1988; Worthy & Beck, 1995).

If discussions are to become a staple feature of classroom instruction, there is need for developments in numerous areas. First, although there is a convergence of theory and data suggesting that high-quality discussions can improve students’ comprehension and learning of text, there is need for more experimental and more rigorous research. As we have indicated, much of the research on classroom discussions has used correlational and single-group pretest-posttest designs. More quasi-experimental and experimental studies of discussion practices, involving rigorous designs, are needed to assess the effects on the quality of classroom discourse as well as on individual student comprehension and learning outcomes. It is especially important to assess students’ comprehension and learning on measures that are independent of the discussion to gauge whether students acquire the habits of mind to transfer their abilities to new texts and novel situations.

Further, research must examine the conduct and effects of discussions outside of literary texts in language arts. Indeed, an emerging trend is that researchers have begun assessing the effects of discussion on content area knowledge when implemented in science (Gillies, Nichols, & Burgh, 2011; Mercer, Dawes, Wegerif, & Sams, 2004; Murphy, 2014; Murphy, Greene, & Butler, 2015) and mathematics (Pauli & Reusser, 2015; Stein, Engle, Smith, & Hughes, 2015) classrooms. Further, researchers are examining the use of discussions as a tool for second language students learning English (Shen, 2013; Zhang, Anderson, & Nguyen-Jahiel, 2013). However, discussions in varying content areas and contexts likely involve a different set of instructional and design
considerations, and more research is needed to examine how these discussions are best conducted and their effects on students’ comprehension and learning.

Second, beyond examining comprehension and learning outcomes as the product of discussions, there is need to examine the dynamics by which students construct mental models of discussions about and around text. As we have indicated, the majority of approaches to classroom discussion attend more to the socially shared and socially constructed nature of learning than to the cognitive internalization of what is being shared. However, as students process discourse, they construct a mental representation in the form of propositions, semantic networks, scripts, or the like. There is need for more microgenetic research that examines how students make use of the outward language tools and signs during discussion and internalize, encode, and process the information in conjunction with the text to construct an elaborated mental representation. As Dewey (1916) and Vygotsky (1978) suggested, the language of the text, the discourse of discussion, as well as the rhetorical skills and dispositions of the discourse community are probably all implicated in the mental representation students construct from discussion. As we outlined earlier, understanding the ways in which classroom discourse enhances students’ comprehension and learning of text will require an integration of the cognitive, social constructivist, and sociocultural theoretical perspectives on discussion.

Third, the notion of traditional classrooms, common since before the times of Aristotle, is quickly evolving as increasing numbers of students are enrolling in online and distance education classrooms. As such, it is of paramount importance that researchers further explore the impact of nontraditional classroom discussions (e.g., online) on students’ learning (Asterhan, 2015; Collins & White, 2015). Indeed, before the aforementioned discussion approaches can be disseminated broadly in large online educational settings, more research is needed to understand the conditions under which these online discussions are most productive (Asterhan, 2015). While this research is still emerging, one exploratory qualitative study assessed the impact of the aforementioned Literature Circles approach in an online setting with groups of graduate students’ Literature Circles (Bowers-Campbell, 2011). After examining the discussion threads, Bowers-Campbell concluded that the “virtual literature circles facilitate[d] socially constructed learning opportunities” (p. 565) and that participants were able to engage in co-construction of meaning as part of their online discussions. Future research must consider the effect of other discussion approaches in online discussions by utilizing experimental or quasi-experimental designs.

Finally, there is need for enhanced professional development to enable teachers to make informed decisions about the use of discussion practices. We have seen in this chapter that there are a plethora of approaches to conducting classroom discussions, that they differ on a number of dimensions relevant to classroom practice, and that the approaches are differentially effective for supporting students’ comprehension and learning. It is important that pre-service and in-service teachers have an in-depth understanding of the similarities and differences among the approaches and their strengths and weaknesses. This will enable teachers to select approaches that are suited to their purposes, to their students, to the subject areas they teach, and to the contexts in which they work. Also important in professional development is for teachers to understand talk and the role of talk in classroom discussions. We know from the research that the success of discussion depends not on simply increasing the amount of student talk, but in enhancing the quality of the talk. We also know, with some degree of reliability, those aspects of discourse and attendant classroom norms that help shape
student comprehension and learning. Teachers need to be familiar with the discourse tools and signs they can use to promote and to recognize productive talk about text and with the kind of classroom culture that is most beneficial for fostering such talk. Through such professional development, teachers will gain an in-depth understanding of how text-based talk can be used as a tool to not only meet their instructional goals, but also to promote thinking and reasoning in their students that transfer to other content areas and contexts—that is, the kind of critical-analytic processing requisite for 21st-century learning.

REFERENCES


Goldenberg (Eds.), *Models of understanding text* (pp. 1–32). Hillsdale, NJ: Lawrence Erlbaum Associates.


